

TRANSPORT DISTANCES AND GEODESIC CONVEXITY FOR SYSTEMS OF DEGENERATE DIFFUSION EQUATIONS

JONATHAN ZINSL AND DANIEL MATTHES

ABSTRACT. We introduce Wasserstein-like dynamical transport distances between vector-valued densities on \mathbb{R} . The mobility function from the scalar theory is replaced by a mobility matrix, that is subject to positivity and concavity conditions. Our primary motivation is to cast certain systems of nonlinear parabolic evolution equations into the variational framework of gradient flows. In the first part of the paper, we investigate the structural properties of the new class of distances like geodesic completeness. The second part is devoted to the identification of λ -geodesically convex functionals and their λ -contractive gradient flows. One of our results is a generalized McCann condition for geodesic convexity of the internal energy. In the third part, the existence of weak solutions to a certain class of degenerate drift-diffusion systems is shown. Even if the underlying energy function is *not* geodesically convex w.r.t. our new distance, the construction of a weak solution is still possible using de Giorgi's *minimizing movement* scheme.

1. INTRODUCTION

1.1. The evolution system and its variational structure. This paper is concerned with the variational structure of the following system of coupled nonlinear evolution equations in one spatial dimension

$$\partial_t \mu(t, x) = \partial_x [\mathbf{M}(\mu(t, x)) \partial_x \mathcal{E}'(\mu(t, x))], \quad t > 0, x \in \mathbb{R}, \quad (1)$$

for the n components μ_1, \dots, μ_n of the sought-for function $\mu : [0, \infty) \times \mathbb{R} \rightarrow S$. We assume that μ attains values in a convex compact set $S \subset \mathbb{R}^n$ with nonempty interior $\text{int}(S)$. Above, $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ is the *mobility matrix*, and \mathcal{E}' is the first variation of the *driving entropy functional* \mathcal{E} which is defined on $\mathcal{M}(\mathbb{R}; S)$, the space of measurable functions on \mathbb{R} with values in S .

Formally, (1) is a gradient flow: solutions $\mu(t, \cdot)$ are curves of steepest descent in the potential landscape of \mathcal{E} , with respect to the Riemannian structure induced on the “manifold” $\mathcal{M}(\mathbb{R}; S)$ by weighted H^{-1} -norms $\|\cdot\|_\mu$ on “tangent vectors” $\dot{\mu}$:

$$\|\dot{\mu}\|_\mu^2 = \int_{\mathbb{R}} \partial_x \Psi^T \mathbf{M}(\mu) \partial_x \Psi \, dx, \quad (2)$$

where the auxiliary function $\Psi : \mathbb{R} \rightarrow \mathbb{R}^n$ solves the elliptic problem

$$\dot{\mu} + \partial_x (\mathbf{M}(\mu) \partial_x \Psi) = 0.$$

This kind of gradient flow structure is well-known in the scalar case $n = 1$, where the mobility matrix \mathbf{M} simplifies to a scalar mobility function \mathbf{m} . In particular, if $\mathbf{m}(z) = z$ is linear, then the metric described above is the L^2 -Wasserstein distance. In the last decade, quite a few well-known evolution equations have been identified as gradient flows in the Wasserstein metric, and have been rigorously analysed on grounds of that special property, among them the Fokker-Planck [17], the porous media [27], the nonlocal aggregation [9], the Hele-Shaw [13] and the fourth order quantum [14] equations. Concerning gradient flows in metrics defined by nonlinear mobility functions \mathbf{m} , we refer to [8, 22].

Comparatively little interest has been devoted to the gradient flow structure of genuine systems (1) with $n > 1$ components. Systems of that kind arise e.g. in reaction-diffusion models for chemical agents as well as for semiconductor dynamics [25, 20, 26, 15], or for population dynamics [10, 19, 18, 5] with or without cross-diffusion.

Date: October 9, 2014.

2010 Mathematics Subject Classification. 49K20, 49J40, 35K40.

Key words and phrases. Dynamical transport distance, metric gradient flow, geodesic convexity, nonlinear mobility.

This research was supported by the German Research Foundation (DFG), Collaborative Research Center SFB-TR 109.

Our main motivation is to analyse systems of the form (1) by means of their variational structure. So far, there does not seem to exist any rigorous study of the properties of the metric induced by (2). In this paper, we provide an extension of several aspects of the rich scalar theory to vector-valued densities μ . Specifically:

- we define a metric $\mathbf{W}_\mathbf{M}$ on $\mathcal{M}(\mathbb{R}; S)$ that gives a rigorous interpretation to the formal structure (2) above, we study the topological properties of the space $(\mathcal{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M})$, and establish connections to the L^2 -Wasserstein distance;
- we derive sufficient conditions for geodesic λ -convexity with respect to $\mathbf{W}_\mathbf{M}$ for a class of entropy functionals \mathcal{E} ;
- we prove existence of weak solutions for a certain class of degenerate drift-diffusion equations of the form (1), even without convexity hypotheses on the respective \mathcal{E} .

We proceed with a summary of our results and relate them to the existing literature.

1.2. Study of the new metric. A cornerstone in the theory of optimal transportation is the Benamou-Brenier dynamical interpretation of the L^2 -Wasserstein distance [4]. Dolbeault *et al.* [12] have used that interpretation to *define* a new class of transportation metrics $\mathbf{W}_\mathbf{m}$, corresponding to *nonlinear* mobilities \mathbf{m} . Their results have been generalized by Lisini and Marigonda [21], and properties of the new metrics have been investigated in [8, 7]. In a nutshell, the Benamou-Brenier formula leads to well-defined metrics $\mathbf{W}_\mathbf{m}$ if the mobility function $\mathbf{m} : S \rightarrow \mathbb{R}$ is positive and concave on the interior of the (possibly semi-infinite) interval $S \subset \mathbb{R}$.

We extend the approach of [12, 21] to densities $\mu : \mathbb{R} \rightarrow S$ with values in a convex and compact set $S \subset \mathbb{R}^n$, and a mobility matrix $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ in place of \mathbf{m} . Our hypotheses on \mathbf{M} are:

- (C0) $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ is continuous, and is smooth on $\text{int}(S)$.
- (C1) $\mathbf{M}(z)$ is symmetric and positive definite for each $z \in \text{int}(S)$.
- (C2) $D^2\mathbf{M}(z)[\zeta, \zeta]$ is negative semidefinite for each $z \in \text{int}(S)$ and $\zeta \in \mathbb{R}^n$.
- (C3) $\mathbf{M}(z)\nu = 0$ if $z \in \partial S$ and ν is a normal vector to ∂S at z .

Conditions (C1)–(C2) are direct generalizations of positivity and concavity of the mobility function \mathbf{m} , and (C0) is a technical hypothesis. Condition (C3) is a natural requirement that is satisfied in all of our examples, but is not substantial for the proofs. Its interpretation is that the values of solutions to (1) are confined to S .

Finally, whenever discussing specific examples, we shall further assume that \mathbf{M} is *induced by a function* $h \in C^2(\text{int}(S))$, which means that

$$\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} \quad \text{for all } z \in \text{int}(S). \quad (3)$$

This hypothesis allows to formulate the multi-component heat equation $\partial_t \mu = \partial_{xx} \mu$ in the form (1); with the functional $\mathcal{E}(\mu) = \int_{\mathbb{R}} h(\mu(x)) dx$.

Under conditions (C0)–(C3), we prove that the Benamou-Brenier formula with the norms from (2) defines a (pseudo-) metric $\mathbf{W}_\mathbf{M}$ on the space $\mathcal{M}(\mathbb{R}; S)$. Moreover, by a careful transfer of the proofs in [12, 21] to the multi-component setting, we obtain that $\mathbf{W}_\mathbf{M}$ inherits the essential topological properties known for the $\mathbf{W}_\mathbf{m}$ distances, like

- existence of constant-speed geodesics connecting densities of finite distance,
- lower semicontinuity with respect to weak* convergence,
- weak*-relative compactness of bounded sets.

We further discuss under which criteria $\mathbf{W}_\mathbf{M}$ is finite.

In practice, the conditions (C0)–(C3) turn out to be quite restrictive, and their validity is fragile under perturbations. A seemingly trivial family of examples is given by the *fully decoupled* mobility matrices,

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & & & \\ & \mathbf{m}_2(z_2) & & \\ & & \ddots & \\ & & & \mathbf{m}_n(z_n) \end{pmatrix}, \quad (4)$$

with n non-negative, concave (scalar) mobility functions $\mathbf{m}_k : [a_k, b_k] \rightarrow \mathbb{R}$. Since the components do not interact with each other through \mathbf{M} , one has that $\mathbf{W}_\mathbf{M}^2 = \mathbf{W}_{\mathbf{m}_1}^2 + \dots + \mathbf{W}_{\mathbf{m}_n}^2$, i.e., $\mathbf{W}_\mathbf{M}$ is simply the sum of the metrics for the components $\mathbf{W}_{\mathbf{m}_k}$. Somewhat surprisingly, it turns out that fully decoupled matrices are ungeneric for property (C2) in the sense that any sufficiently general, arbitrarily small perturbation of the

components of \mathbf{M} makes (C2) invalid. We shall show how certain fully decoupled matrices can be “stabilized” with a suitably chosen special perturbation such that the perturbed mobility matrix retains (C2).

1.3. Geodesic convexity. Geodesic λ -convexity plays a pivotal role in the theory of metric gradient flows [3]. For the gradient flows of λ -convex functionals, one obtains immediately contractivity estimates, asymptotics for the long-time behaviour, and error bounds on the time-discrete implicit Euler approximation. However, λ -convexity in transportation metrics is a very rare property [24].

Up to now, the only known λ -convex functionals \mathcal{E} for the metrics $\mathbf{W}_{\mathbf{m}}$ with nonlinear mobilities \mathbf{m} in space dimension $d = 1$ are the *internal energies*

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\mu(x)) \, dx, \quad (5)$$

provided that f satisfies the generalized McCann condition [8] in $d = 1$, i.e., $s \mapsto \mathbf{m}(s)^2 f''(s)$ is a nonnegative function, and the *regularized potential energies*

$$\mathcal{V}(\mu) = \int_{\mathbb{R}} [\alpha h(\mu(x)) + \rho(x)\mu(x)] \, dx, \quad (6)$$

where $h : [0, \infty) \rightarrow \mathbb{R}$ is such that $\mathbf{m}h'' \equiv 1$, $\alpha > 0$, and $\rho : \mathbb{R} \rightarrow \mathbb{R}$ is a smooth function of compact support [22]. The respective gradient flows are given by

$$\partial_t \mu = \partial_{xx} P(\mu), \quad \text{and} \quad \partial_t \mu = \alpha \partial_{xx} \mu + \partial_x (\mathbf{m}(\mu) \partial_x \rho),$$

where $P'(z) = \mathbf{m}(z)f''(z)$.

Both types of functionals (5) and (6) possess canonical generalizations to densities with multiple components. In (5), simply replace f by a smooth function $f : S \rightarrow \mathbb{R}$. To make sense of (6), assume that \mathbf{M} is induced by $h : S \rightarrow \mathbb{R}$, see (3), and use a potential $\rho : \mathbb{R} \rightarrow \mathbb{R}^n$ with n components. We derive sufficient criteria for the geodesic λ -convexity of these functionals with respect to the new metric $\mathbf{W}_{\mathbf{M}}$. For that, we use the formalism developed by Liero and Mielke [25, 20], which is based on the *Eulerian calculus* for transportation distances, see [11, 28].

Our own generalization of McCann’s condition for \mathcal{E} of the form (5) is given in Proposition 5.3, see formula (44). Our examples for pairs of a (nondiagonal) mobility matrix \mathbf{M} and a function f that satisfy this condition are currently limited to perturbations of the heat equation. For definiteness, assume that \mathbf{M} is induced by h and choose $f = h + \varepsilon g$, where $g : S \rightarrow \mathbb{R}$ vanishes near the boundary of S . If ε is sufficiently small, then our generalized McCann condition holds. The evolution equation (1) specializes to a perturbation of the multi-component heat equation:

$$\partial_t \mu = \partial_{xx} \mu + \varepsilon \partial_x (\mathbf{M}(\mu) \partial_x (\nabla_z g(\mu))).$$

In contrast, if \mathbf{M} is a fully decoupled mobility, \mathcal{E} has to be decoupled in order to satisfy our generalized McCann condition.

Our condition assuring λ -convexity for functionals of type (6) is given in (47). Even for smooth ρ of compact support, it imposes an apparently very strong restriction on the function h in $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1}$.

1.4. Construction of solutions to (1). In Section 6, we discuss the primary application of the new metric $\mathbf{W}_{\mathbf{M}}$, namely the construction of weak solutions to a class of drift-diffusion equations of the form (1) by means of de Giorgi’s *minimizing movement scheme*, which is a time-discrete implicit Euler scheme for gradient flows (cf. [17], see also Section 2 below). Specifically, we consider the initial value problem

$$\partial_t \mu = \partial_x (\mathbf{M}(\mu) \nabla_z^2 f(\mu) \partial_x \mu + \mathbf{M}(\mu) \partial_x \eta), \quad \mu(0) = \mu^0, \quad (7)$$

where the mobility matrix \mathbf{M} is fully decoupled as in (4), $S \subset \mathbb{R}^n$ is an n -cuboid and $f : S \rightarrow \mathbb{R}$ is uniformly convex, $\nabla_z^2 f(z) \geq C_f \mathbf{1}$ with $C_f > 0$, but does *not* need to be the sum of functions of the components of μ . Thus, the diffusion matrix $\mathbf{M} \nabla_z^2 f$ will not be symmetric nor positive definite in general. Also, the corresponding energy functional

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\mu(x)) \, dx + \int_{\mathbb{R}} \mu(x)^T \eta(x) \, dx$$

will not be λ -convex.

Still, the variational minimizing movement scheme is well-posed. We prove that in the limit of vanishing time step size, it produces a limit curve that is a weak solution to (7), see Theorem 6.10. The crucial *a priori*

estimate for the passage to the limit is provided by the dissipation of the auxiliary functional $\int_{\mathbb{R}} h(\mu(x)) dx$, where h induces \mathbf{M} according to (3). That dissipation amounts to

$$-\frac{d}{dt} \int_{\mathbb{R}} h(\mu) dx \geq C_f \int_{\mathbb{R}} |\partial_x \mu|^2 dx + \int_{\mathbb{R}} \partial_x \eta^T \partial_x \mu dx, \quad (8)$$

and therefore provides square-integrability of $\partial_x \mu$ in space and time.

We emphasize that the global existence of solutions to (7) is a nontrivial result of independent interest. It does not follow immediately from classical parabolic theory: Indeed, since $\mathbf{M} \nabla_z^2 f$ typically lacks positivity (meaning that $v^T \mathbf{M} \nabla_z^2 f v \geq 0$), the differential operator in (7) is not elliptic in the strong sense. The theory for parabolic equations with normally elliptic operators, see e.g. [1], provides existence of solutions only locally in time for sufficiently regular initial data; for extension of those to global solutions, additional estimates would be needed which guarantee that the values of the solution μ stay away from the boundary of the admissible set S .

Our gradient flow approach is conceptually different. The minimizing movement scheme naturally produces a globally defined curve μ with values in S in the continuous time limit. Instead, the main step of the proof is to identify this limit curve as a weak solution to (7), using the compactness induced by (8). In comparison to the classical results, we obtain weaker solutions of lower regularity, but we can allow for more general initial data.

We remark that just recently, a closely related class of *reaction-diffusion* systems has been studied [18] on grounds of a very similar dissipation estimate as in (8), see also [6]. However, the equations considered in [18] are generically not of gradient flow type, and a completely different technique, based on a suitable transformation of the dependent variables, has been employed for approximation of global weak solutions with values in a prescribed set S . The connection to our treatment of (7) is that in both cases, subtle structural properties of the diffusion matrix play a pivotal role.

1.5. Plan of the paper. In Section 2, we introduce basic notation and definitions. Then we provide a couple of examples for mobility matrices \mathbf{M} satisfying (C0)–(C3) in Section 3. In Section 4, the general theory for distances generated by a mobility matrix is developed. There, we start with basic properties of the objects occurring in the definition of these new distances (Section 4.1) before investigating solutions to the continuity equation (Section 4.2). This enables us to prove the defining properties of the distance and several additional topological properties in Section 4.3.

Section 5 is devoted to geodesic convexity and gradient flows with respect to the distances established in Section 4 and begins with an introduction of the abstract background (Section 5.1). We continue with the investigation of the multi-component heat equation (Section 5.2) and general internal energy functionals (Section 5.3), where specific perturbation results are also given. Afterwards, we study geodesic convexity of the regularized potential energy functionals in Section 5.4.

In Section 6, we first introduce the framework (Section 6.1) before constructing an approximate time-discrete solution to the given equation in Section 6.2. The desired weak solution is then obtained by passage to the continuous-time limit (Section 6.3).

2. PRELIMINARIES

We first introduce our notation before stating relevant definitions and statements and refer to [2, 30, 3] for more details on optimal transportation, gradient flows, and their measure theoretic preliminaries.

2.1. Basic notation. Components of a n -vector $v \in \mathbb{R}^n$ are indicated with lower indices: $v = (v_1, v_2, \dots, v_n)$. By $|\cdot|$, we denote the usual Euclidean norm and inner product on \mathbb{R}^n , whereas $\langle \cdot, \cdot \rangle$ formally denotes the duality pairing on $L^2(\mathbb{R}; \mathbb{R}^n)$. Inequalities between vectors, multi-dimensional intervals (also referred to as n -cuboids) $[q_0, q_1]$ for $q_0, q_1 \in \mathbb{R}^n$, $q_0 \leq q_1$, as well as integration of vector-valued functions are understood component-wise.

We use ∇_z for the gradient, ∇_z^2 for the Hessian and D in combination with square brackets for directional derivatives with respect to z . For instance, if $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ and $\mu : \mathbb{R} \rightarrow S$, then we write the chain rule as

$$\partial_x \mathbf{M}(\mu) = D\mathbf{M}(\mu)[\partial_x \mu].$$

Note that even for symmetric matrices \mathbf{M} , the third order-tensor \mathbf{DM} and the fourth-order tensor $\mathbf{D}^2\mathbf{M}$ are not totally symmetric in general, although $\mathbf{DM}(z)[\zeta]$ and $\mathbf{D}^2\mathbf{M}(z)[\zeta, \tilde{\zeta}]$ are symmetric $n \times n$ matrices, for every choice of $\zeta, \tilde{\zeta} \in \mathbb{R}^n$. Given a multilinear operator or its tensor representative, the norm $\|\cdot\|$ denotes the operator norm.

For a nonnegative measurable function $\tilde{\mu} : \mathbb{R} \rightarrow \mathbb{R}^n$, the functional

$$\ell_2(\tilde{\mu}) := \int_{\mathbb{R}} x^2 \mathbf{e}^T \tilde{\mu}(x) dx \in \mathbb{R} \cup \{\infty\}$$

is called the *second moment* of $\tilde{\mu}$, where $\mathbf{e} := (1, 1, \dots, 1)^T \in \mathbb{R}^n$. The space of nonnegative functions $\nu \in L^1(\mathbb{R})$ with fixed mass and finite second moment can be equipped with the L^2 -Wasserstein distance \mathbf{W}_2 :

$$\mathbf{W}_2(\nu, \tilde{\nu}) := \inf \left[\int_{\mathbb{R} \times \mathbb{R}} |x - y|^2 d\boldsymbol{\nu} : \boldsymbol{\nu} \in \Gamma(\nu, \tilde{\nu}) \right]^{1/2},$$

where $\Gamma(\nu, \tilde{\nu})$ denotes the set of all couplings between the two finite Borel measures with density ν and $\tilde{\nu}$, respectively. By [4], one has the equivalent dynamic characterization of the Wasserstein distance:

$$\mathbf{W}_2(\nu, \tilde{\nu}) = \inf \left[\int_0^1 \int_{\mathbb{R}} \dot{\nu}(t, x) |v(t, x)|^2 dx dt : \partial_t \hat{\nu} + \partial_x(\hat{\nu} v) = 0 \text{ in } (0, 1) \times \mathbb{R}, \hat{\nu}|_{t=0} = \nu, \hat{\nu}|_{t=1} = \tilde{\nu} \right]^{1/2}.$$

Given a closed set $A \subset \mathbb{R}^n$, $\mathcal{M}(\mathbb{R}; A)$ denotes the space of all measurable functions $\tilde{\mu} : \mathbb{R} \rightarrow A$. We call a sequence of measurable functions $(\tilde{\mu}_k)_{k \in \mathbb{N}}$ in $\mathcal{M}(\mathbb{R}; A)$ weak*-convergent to its limit $\tilde{\mu} \in \mathcal{M}(\mathbb{R}; A)$, if for all $\rho \in C_c^0(\mathbb{R}; \mathbb{R}^n)$, one has

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}} \tilde{\mu}_k^T \rho dx = \int_{\mathbb{R}} \tilde{\mu}^T \rho dx.$$

2.2. Gradient flows in metric spaces. For general metric spaces (\mathbf{X}, \mathbf{d}) , a functional $\mathcal{A} : \mathbf{X} \rightarrow \mathbb{R} \cup \{\infty\}$ is called λ -geodesically convex w.r.t. \mathbf{d} for some $\lambda \in \mathbb{R}$, if for all $u_0, u_1 \in \mathcal{A}$ and all $t \in [0, 1]$, one has

$$\mathcal{A}(u_t) \leq (1 - t)\mathcal{A}(u_0) + t\mathcal{A}(u_1) - \frac{\lambda}{2}t(1 - t)\mathbf{d}^2(u_0, u_1),$$

where $u_t : [0, 1] \rightarrow \mathbf{X}$ is a geodesic curve connecting u_0 and u_1 . We introduce the notion of λ -contractive gradient flow by means of the following

Definition 2.1 (λ -flow). *Let $\mathcal{A} : \mathbf{X} \rightarrow \mathbb{R} \cup \{\infty\}$ be a lower semicontinuous functional on the metric space (\mathbf{X}, \mathbf{d}) . A continuous semigroup \mathbf{S} on (\mathbf{X}, \mathbf{d}) is called λ -flow for some $\lambda \in \mathbb{R}$, if the evolution variational estimate (with parameter λ)*

$$\frac{1}{2} \frac{d^+}{dt} \mathbf{d}^2(\mathbf{S}^t(u), \tilde{u}) + \frac{\lambda}{2} \mathbf{d}^2(\mathbf{S}^t(u), \tilde{u}) + \mathcal{A}(\mathbf{S}^t(u)) \leq \mathcal{A}(\tilde{u}), \quad (9)$$

holds for arbitrary u, \tilde{u} in the domain of \mathcal{A} , and for all $t \geq 0$.

By [11], λ -geodesic convexity is implied by (9). Henceforth, \mathbf{S} is called *gradient flow* of \mathcal{A} with respect to the distance \mathbf{d} .

A possible method to construct a gradient flow is by means of the so-called *minimizing movement scheme*.

Definition 2.2 (Minimizing movement). *Given a step size $\tau > 0$ and an initial value u_τ^0 , determine u_τ^k inductively for $k \in \mathbb{N}$ as minimizers of*

$$\mathcal{A}_\tau(u | u_\tau^{k-1}) := \frac{1}{2\tau} \mathbf{d}^2(u, u_\tau^{k-1}) + \mathcal{A}(u), \quad (10)$$

which exist under suitable conditions on the functional \mathcal{A} . Then define a time-discrete solution $u_\tau : [0, \infty) \rightarrow \mathbf{X}$ by piecewise constant interpolation:

$$u_\tau(t) := u_\tau^k \text{ for } t \in ((k-1)\tau, k\tau]. \quad (11)$$

Theorem 2.3 (Flow interchange lemma [23, Thm. 3.2]). *Let \mathcal{B} be a proper, lower semicontinuous and λ -geodesically convex functional on (\mathbf{X}, \mathbf{d}) . Let furthermore \mathcal{A} be another proper, lower semicontinuous functional on (\mathbf{X}, \mathbf{d}) such that $\text{Dom}(\mathcal{A}) \subset \text{Dom}(\mathcal{B})$. Assume that, for arbitrary $\tau > 0$ and $\tilde{u} \in \mathbf{X}$, the functional $\mathcal{A}_\tau(\cdot | \tilde{u})$ possesses a minimizer u . Then, the following holds:*

$$\mathcal{B}(u) + \tau D^{\mathcal{B}} \mathcal{A}(u) + \frac{\lambda}{2} \mathbf{d}^2(u, \tilde{u}) \leq \mathcal{B}(\tilde{u}).$$

There, $D^{\mathcal{B}} \mathcal{A}(u)$ denotes the dissipation of the functional \mathcal{A} along the gradient flow $S_{\mathcal{B}}^{(\cdot)}$ of the functional \mathcal{B} , i.e.

$$D^{\mathcal{B}} \mathcal{A}(u) := \limsup_{h \searrow 0} \frac{\mathcal{A}(u) - \mathcal{A}(S_{\mathcal{B}}^h(u))}{h}.$$

In order to define a time-continuous flow, passage to the limit $\tau \searrow 0$ is necessary. The following theorem provides a useful tool in this context.

Theorem 2.4 (Extension of the Aubin-Lions lemma [29, Thm. 2]). *Let \mathbf{Y} be a Banach space and $\mathcal{A} : \mathbf{Y} \rightarrow [0, \infty]$ be lower semicontinuous and have relatively compact sublevels in \mathbf{Y} . Let furthermore $\mathbf{W} : \mathbf{Y} \times \mathbf{Y} \rightarrow [0, \infty]$ be lower semicontinuous and such that $\mathbf{W}(u, \tilde{u}) = 0$ for $u, \tilde{u} \in \text{Dom}(\mathcal{A})$ implies $u = \tilde{u}$.*

Let $(U_k)_{k \in \mathbb{N}}$ be a sequence of measurable functions $U_k : (0, T) \rightarrow \mathbf{Y}$. If

$$\sup_{k \in \mathbb{N}} \int_0^T \mathcal{A}(U_k(t)) dt < \infty, \quad (12)$$

$$\limsup_{h \searrow 0} \sup_{k \in \mathbb{N}} \mathbf{W}(U_k(t+h), U_k(t)) dt = 0, \quad (13)$$

then there exists a subsequence that converges in measure w.r.t. $t \in (0, T)$ to a limit $U : (0, T) \rightarrow \mathbf{Y}$.

3. EXAMPLES OF MOBILITY MATRICES

This section is devoted to examples of mobility matrices $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ that satisfy conditions (C0)–(C3) stated in the introduction. We will occasionally also consider the following stronger version of (C2):

(C2') The matrix $D^2 \mathbf{M}(z)[\zeta, \zeta] \in \mathbb{R}^{n \times n}$ is negative definite for all $z \in \text{int}(S)$ and $\zeta \in \mathbb{R}^n \setminus \{0\}$.

All of our examples are of the form (3), where \mathbf{M} is induced by a convex function h .

3.1. Fully decoupled mobilities. Consider concave functions $\mathbf{m}_1, \dots, \mathbf{m}_n$ with $\mathbf{m}_j : [S_j^\ell, S_j^r] \rightarrow \mathbb{R}$, $S_j^\ell < S_j^r$, such that $\mathbf{m}_j(s) > 0$ for $s \in (S_j^\ell, S_j^r)$ and $\mathbf{m}_j(S_j^\ell) = \mathbf{m}_j(S_j^r) = 0$, for each j . Define a mobility matrix $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ on the n -cuboid $S := [S^\ell, S^r]$ by

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & & \\ & \ddots & \\ & & \mathbf{m}_n(z_n) \end{pmatrix}. \quad (14)$$

Clearly \mathbf{M} is of the form (3), where

$$h(z) = h_1(z_1) + \dots + h_n(z_n),$$

and each $h_j : (S_j^\ell, S_j^r) \rightarrow \mathbb{R}$ is a second primitive of the respective $\frac{1}{\mathbf{m}_j}$, i.e., $\mathbf{m}_j(s)h_j''(s) = 1$. It is immediately verified that \mathbf{M} satisfies (C0)–(C3). Concerning property (C2), we remark that

$$\beta^T D^2 \mathbf{M}(z)[\zeta, \zeta] \beta = \sum_{j=1}^n \mathbf{m}_j''(z) (\zeta_j \beta_j)^2,$$

hence the sharper condition (C2') is *not* satisfied, even if all \mathbf{m}_j are *strictly* concave functions. This is the reason why the concavity (C2) is lost under generic perturbations of \mathbf{M} . In the next example below, we discuss a very special “perturbation” of a particular matrix of type (14), for which (C2') is valid.

For obvious reasons, we call mobility matrices \mathbf{M} of the form (14) *fully decoupled*: the different species do not influence each other’s mobility. It is clear that each fully decoupled matrix \mathbf{M} induces a metric on $\mathcal{M}(\mathbb{R}; S)$, simply applying the theory from [12, 21] to each component separately.

3.2. Perturbations of a fully decoupled mobility. Let us now specialize the previous example by choosing $n = 2$ components, $S = [0, 1]^2$ and $h_0 : (0, 1)^2 \rightarrow \mathbb{R}$ with

$$h_0(z) = z_1 \log z_1 + (1 - z_1) \log(1 - z_1) + z_2 \log z_2 + (1 - z_2) \log(1 - z_2). \quad (15)$$

From (3), we obtain the fully decoupled mobility matrix

$$\mathbf{M}_0(z) = (\nabla_z^2 h_0(z))^{-1} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}, \quad \text{with } d_j = \mathbf{m}(z_j),$$

where $\mathbf{m}(s) = s(1 - s)$. By the discussion above, (C0)–(C3) are satisfied, but (C2') is not. It is easily seen that for a general (smooth, compactly supported) function $g : (0, 1)^2 \rightarrow \mathbb{R}$, the matrix $\tilde{\mathbf{M}}_\varepsilon = (\nabla_z^2(h_0 + \varepsilon g))^{-1}$ does *not* satisfy (C2) anymore, no matter how small $\varepsilon > 0$ is.

Let us introduce a very special perturbation h_ε of h_0 :

$$h_\varepsilon(z) = h_0(z) + \varepsilon z_1 z_2 (1 - z_1)(1 - z_2) = h_0(z) + \varepsilon d_1 d_2. \quad (16)$$

We are going to show that $\mathbf{M}_\varepsilon(z) = (\mathbf{H}_\varepsilon(z))^{-1}$, with

$$\mathbf{H}_\varepsilon(z) := \nabla_z^2 h_\varepsilon(z) = \begin{pmatrix} \frac{1}{d_1} & 0 \\ 0 & \frac{1}{d_2} \end{pmatrix} + \varepsilon \begin{pmatrix} -2d_2 & d'_1 d'_2 \\ d'_1 d'_2 & -2d_1 \end{pmatrix}, \quad \text{with } d'_1 = 1 - 2z_1, \ d'_2 = 1 - 2z_2,$$

satisfies (C0)–(C3), and in addition also (C2'), for all sufficiently small $\varepsilon > 0$. Thus, this special perturbation makes the mobility matrix robust with respect to further (smaller) generic perturbations.

First, note that $\mathbf{M}_\varepsilon(z)$ is well-defined at $z \in (0, 1)^2$ if

$$\det \mathbf{H}_\varepsilon(z) = \frac{1}{d_1 d_2} - 4\varepsilon + \varepsilon^2 (4d_1 d_2 - (d'_1 d'_2)^2) \quad (17)$$

is positive. This is true simultaneously at all $z \in (0, 1)^2$ if $\varepsilon > 0$ is sufficiently small. It is further easily seen that \mathbf{M}_ε extends continuously to the boundary of S by setting $\mathbf{M}_\varepsilon(z) = \mathbf{M}_0(z)$ for $z \in \partial S$; just observe that

$$\mathbf{M}_\varepsilon(z) = \frac{1}{1 - \varepsilon d_1 d_2 [4 - \varepsilon (4d_1 d_2 - (d'_1 d'_2)^2)]} \left[\mathbf{M}_0(z) - \varepsilon d_1 d_2 \begin{pmatrix} 2d_1 & d'_1 d'_2 \\ d'_1 d'_2 & 2d_2 \end{pmatrix} \right],$$

and that $d_1 d_2 \searrow 0$ as $z \rightarrow \partial S$. This implies (C0) and (C3) for \mathbf{M}_ε .

Next, since the entries of \mathbf{M}_ε vary continuously with ε , and since $\det \mathbf{M}_\varepsilon(z) = (\det \mathbf{H}_\varepsilon(z))^{-1}$ never vanishes for any $z \in (0, 1)^2$ and any sufficiently small $\varepsilon > 0$, it follows that \mathbf{M}_ε inherits the positive definiteness of \mathbf{M}_0 . Thus, also (C1) is verified.

The proof of condition (C2') is more involved. To begin with, observe that $\mathbf{M}_\varepsilon(z) = (\mathbf{H}_\varepsilon(z))^{-1}$ implies

$$\mathbf{D}^2 \mathbf{M}_\varepsilon(z)[\zeta, \zeta] = -\mathbf{H}_\varepsilon(z)^{-1} \mathbf{D}^2 \mathbf{H}_\varepsilon(z)[\zeta, \zeta] \mathbf{H}_\varepsilon(z)^{-1} + 2\mathbf{H}_\varepsilon(z)^{-1} \mathbf{D} \mathbf{H}_\varepsilon(z)[\zeta] \mathbf{H}_\varepsilon(z)^{-1} \mathbf{D} \mathbf{H}_\varepsilon(z)[\zeta] \mathbf{H}_\varepsilon(z)^{-1}.$$

Thus, for proving (C2'), it suffices to show that for all $z \in (0, 1)^2$ and all $\beta, \zeta \in \mathbb{R}^n \setminus \{0\}$,

$$P := (d_1 d_2)^3 \beta^T (\det \mathbf{H}_\varepsilon(z) \mathbf{D}^2 \mathbf{H}_\varepsilon(z)[\zeta, \zeta] - 2\mathbf{D} \mathbf{H}_\varepsilon(z)[\zeta] \mathbf{H}_\varepsilon(z)^+ \mathbf{D} \mathbf{H}_\varepsilon(z)[\zeta]) \beta > 0,$$

where $\det \mathbf{H}_\varepsilon$ is given in (17), and

$$\begin{aligned} \mathbf{D} \mathbf{H}_\varepsilon(z)[\zeta] &= \begin{pmatrix} -\frac{d'_1}{d_1^2} \zeta_1 & 0 \\ 0 & -\frac{d'_2}{d_2^2} \zeta_2 \end{pmatrix} - 2\varepsilon \begin{pmatrix} d'_2 \zeta_2 & d'_1 \zeta_1 + d'_2 \zeta_1 \\ d'_1 \zeta_2 + d'_2 \zeta_1 & d'_1 \zeta_1 \end{pmatrix}, \\ \mathbf{D}^2 \mathbf{H}_\varepsilon(z)[\zeta, \zeta] &= 2 \begin{pmatrix} \frac{1-3d_1}{d_1^3} \zeta_1^2 & 0 \\ 0 & \frac{1-3d_2}{d_2^3} \zeta_2^2 \end{pmatrix} + 4\varepsilon \begin{pmatrix} \zeta_2^2 & 2\zeta_1 \zeta_2 \\ 2\zeta_1 \zeta_2 & \zeta_1^2 \end{pmatrix}, \\ \mathbf{H}_\varepsilon(z)^+ &= \det \mathbf{H}_\varepsilon(z) \mathbf{H}_\varepsilon(z)^{-1} = \begin{pmatrix} \frac{1}{d_2} & 0 \\ 0 & \frac{1}{d_1} \end{pmatrix} + \varepsilon \begin{pmatrix} -2d_1 & -d'_1 d'_2 \\ -d'_1 d'_2 & -2d_2 \end{pmatrix}. \end{aligned}$$

A tedious but straightforward calculation leads to the following explicit representation of P , with the abbreviations $\tilde{\zeta}_1 := d_2 \zeta_1$, $\tilde{\zeta}_2 := d_1 \zeta_2$:

$$\begin{aligned} P &= 2[\tilde{\zeta}_1^2 + 2\varepsilon(d_2 \tilde{\zeta}_2)^2] \beta_1^2 + 2[\tilde{\zeta}_2^2 + 2\varepsilon(d_1 \tilde{\zeta}_1)^2] \beta_2^2 + \varepsilon[\hat{f}_1 \tilde{\zeta}_1^2 + \varepsilon f_2(d_2 \tilde{\zeta}_2)^2 + \hat{f}_3 \tilde{\zeta}_1^1(d_2 \tilde{\zeta}_2)] \beta_1^2 \\ &\quad + \varepsilon[\check{f}_1 \tilde{\zeta}_2^2 + \varepsilon f_2(d_1 \tilde{\zeta}_1)^2 + \check{f}_3 \tilde{\zeta}_2(d_1 \tilde{\zeta}_1)] \beta_2^2 + \varepsilon[\hat{f}_4 \tilde{\zeta}_1(d_1 \tilde{\zeta}_1) + \check{f}_4 \tilde{\zeta}_2(d_2 \tilde{\zeta}_2) + 2f_5 \tilde{\zeta}_1 \tilde{\zeta}_2] \beta_1 \beta_2, \end{aligned}$$

where the functions f_i , \hat{f}_i and \check{f}_i are bounded, uniformly with respect to $z \in (0, 1)^2$ and (small) $\varepsilon > 0$:

$$\begin{aligned}\hat{f}_1 &:= 2d_2(8d_1 - 3) + \varepsilon(d_2(-64d_1 + 132d_1^2 + 8) - 39d_1^2 - 2 + 18d_1) - 8\varepsilon^2 d_1^3 d_2^2(1 - 4d_2), \\ \check{f}_1 &:= 2d_1(8d_2 - 3) + \varepsilon(d_1(-64d_2 + 132d_2^2 + 8) - 39d_2^2 - 2 + 18d_2) - 8\varepsilon^2 d_2^3 d_1^2(1 - 4d_1), \\ f_2 &:= 4d_1 d_2 - d_1 - d_2 - \varepsilon d_1 d_2(28d_1 d_2 - 10d_1 - 10d_2 + 3), \\ \hat{f}_3 &:= 4d_1' d_2'(1 + \varepsilon d_2(1 - 2\varepsilon(1 - 2d_2)d_1^2)), \\ \check{f}_3 &:= 4d_1' d_2'(1 + \varepsilon d_1(1 - 2\varepsilon(1 - 2d_1)d_2^2)), \\ \hat{f}_4 &:= -2d_1' d_2'(1 + 4\varepsilon(d_1 - 1 + 2\varepsilon d_2 d_1^2(1 - 2d_2))), \\ \check{f}_4 &:= -2d_1' d_2'(1 + 4\varepsilon(d_2 - 1 + 2\varepsilon d_1 d_2^2(1 - 2d_1))), \\ f_5 &:= 1 + 40d_1 d_2 - 6d_1 - 6d_2 - 32\varepsilon d_1^2 d_2^2 - 4\varepsilon^2 d_1^2 d_2^2(5 + 56d_1 d_2 - 18d_1 - 18d_2).\end{aligned}$$

From elementary calculations – applying the Cauchy-Schwarz inequality and collecting terms – we conclude that

$$P \geq [\tilde{\zeta}_1^2 + 2\varepsilon(d_2 \tilde{\zeta}_2)^2] \beta_1^2 + [\tilde{\zeta}_2^2 + 2\varepsilon(d_1 \tilde{\zeta}_1)^2] \beta_2^2$$

for arbitrary $z \in (0, 1)^2$, $\beta, \zeta \in \mathbb{R}^n$, and all sufficiently small $\varepsilon > 0$. This implies positivity of P for $\beta \neq 0$ and $\zeta \neq 0$, and therefore proves (C2').

3.3. Volume filling mobility. The following example describes the interaction of species that influence each other's mobilities by competing for limited volume. This example is related but not identical to the one considered in [20], where in addition a microlocal conservation of mass was assumed.

Define the n^{th} standard simplex

$$S := \left\{ z \in [0, 1]^n : 1 - \sum_{j=1}^n z_j \geq 0 \right\}$$

as state space and the map $h : \text{int}(S) \rightarrow \mathbb{R}$ by

$$h(z) := \sum_{j=1}^n z_j \log z_j + \left(1 - \sum_{j=1}^n z_j \right) \log \left(1 - \sum_{j=1}^n z_j \right).$$

The second order derivatives of h amount to

$$\frac{\partial^2 h}{\partial z_i \partial z_j}(z) = \frac{1}{z_i} \delta_{ij} + \frac{1}{1 - \sum_{\ell=1}^n z_\ell},$$

where δ_{ij} denotes Kronecker's delta. By elementary calculations, we obtain the explicit form of the inverse matrix,

$$\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix} - z z^T.$$

Property (C0) obviously holds. To verify (C1), let $\gamma \in \mathbb{R}^n$ be given and observe that

$$\gamma^T z z^T \gamma = \sum_{i,j=1}^n z_i z_j \gamma_i \gamma_j \leq \frac{1}{2} \sum_{i,j=1}^n z_i z_j (\gamma_i^2 + \gamma_j^2) = \left(\sum_{j=1}^n \gamma_j^2 z_j \right) \left(\sum_{\ell=1}^n z_\ell \right).$$

Therefore,

$$\gamma^T \mathbf{M}(z) \gamma = \sum_{j=1}^n \gamma_j^2 z_j - \gamma^T z z^T \gamma \geq \sum_{j=1}^k \gamma_j^2 z_j \left(1 - \sum_{\ell=1}^n z_\ell \right),$$

which is positive for all $z \in \text{int}(S)$ and $\gamma \neq 0$. Condition (C2) is immediately obtained from

$$D^2 \mathbf{M}(z)[\zeta, \zeta] = -2\zeta \zeta^T,$$

which is negative semidefinite, for arbitrary $z \in \text{int}(S)$ and $\zeta \in \mathbb{R}^n$. Note that $D^2\mathbf{M}(z)[\zeta, \zeta]$ has rank one, hence the stronger condition (C2') is not satisfied. Finally, let $z \in \partial S$, and let ν be a normal vector to ∂S at z . We distinguish two cases. In the first, z lies on one of the coordinate hyperplanes. Then $\nu_j \neq 0$ only if $z_j = 0$, for $j = 1, \dots, n$, and so clearly $\mathbf{M}(z)\nu = 0$. In the second case, we have $z_1 + \dots + z_n = 1$. Hence the normal vector is (a multiple of) $\mathbf{e} = (1, \dots, 1)^T$, and therefore

$$\mathbf{M}(z)\mathbf{e} = z - z(z^T\mathbf{e}) = \left(1 - \sum_{\ell=1}^n z_\ell\right) z = 0.$$

This proves (C3).

3.4. Radially symmetric mobility. On the n -dimensional closed unit ball $S := \overline{\mathbb{B}_1(0)}$, define $h : S \rightarrow \mathbb{R}$ by

$$h(z) = \log(1 + \sqrt{1 - |z|^2}) - \sqrt{1 - |z|^2}.$$

One easily verifies that

$$\nabla_z^2 h(z) = \frac{1}{1 + \sqrt{1 - |z|^2}} \mathbb{1} + \frac{1}{(1 + \sqrt{1 - |z|^2})\sqrt{1 - |z|^2}} \frac{zz^T}{|z|^2},$$

which obviously is positive definite for $z \in \mathbb{B}_1(0)$. Now define \mathbf{M} by (3), i.e.,

$$\begin{aligned} \mathbf{M}(z) &= (\nabla_z^2 h(z))^{-1} \\ &= (1 + \sqrt{1 - |z|^2}) \mathbb{1} + [(1 + \sqrt{1 - |z|^2})\sqrt{1 - |z|^2} - (1 + \sqrt{1 - |z|^2})] \frac{zz^T}{|z|^2} \\ &= (1 + \sqrt{1 - |z|^2}) \mathbb{1} - zz^T. \end{aligned}$$

Conditions (C0) and (C1) obviously hold. Next, for arbitrary $\zeta \in \mathbb{R}^n$ and $z \in \text{int}(S)$, we have that

$$\begin{aligned} D\mathbf{M}(z)[\zeta] &= -(1 - |z|^2)^{-1/2}(z^T\zeta)\mathbb{1} - z\zeta^T - \zeta z^T \\ D^2\mathbf{M}(z)[\zeta, \zeta] &= -(1 - |z|^2)^{-3/2}(z^T\zeta)^2\mathbb{1} - 2\zeta\zeta^T - (1 - |z|^2)^{-1/2}|\zeta|^2\mathbb{1}. \end{aligned}$$

The last matrix is obviously negative definite for each $z \in \mathbb{B}_1(0)$ and $\zeta \neq 0$, which shows (C2'). Finally, to verify (C3), let $z \in S$ with $|z| = 1$ be given, and observe that z itself is a normal vector to ∂S at z . One has

$$\mathbf{M}(z)z = (1 + \sqrt{1 - |z|^2})z - |z|^2z = z - z = 0.$$

4. DISTANCES GENERATED BY A MOBILITY MATRIX

This section is devoted to the study of transport distances between vector-valued densities on \mathbb{R} . Throughout this section, let some convex and compact set $S \subset \mathbb{R}^n$ with nonempty interior be fixed, and recall that $\mathcal{M}(\mathbb{R}; S)$ is the space of measurable functions on \mathbb{R} with values in S . Throughout this section, we assume that $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ is a mobility matrix that satisfies (C0)–(C3).

The object of central interest in this section is the function $\mathbf{W}_{\mathbf{M}} : \mathcal{M}(\mathbb{R}; S) \times \mathcal{M}(\mathbb{R}; S) \rightarrow [0, \infty]$, defined by

$$\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) := \left[\inf \left\{ \int_0^1 \int_{\mathbb{R}} w^T (\mathbf{M}(\mu))^{-1} w \, dx \, dt : (\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1) \right\} \right]^{1/2}, \quad (18)$$

where $\mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ denotes the set of all curves $(\mu, w) = (\mu_t, w_t)_{t \in [0,1]}$ satisfying the *continuity equation*

$$\partial_t \mu + \partial_x w = 0, \quad (19)$$

having μ_0 and μ_1 as starting and terminal values, respectively. We begin by giving a rigorous definition of the objects occurring above.

4.1. The action density function.

Proposition 4.1 (Properties of the density function). *The action density function $\tilde{\phi} : \text{int}(S) \times \mathbb{R}^n \rightarrow [0, \infty)$, defined by*

$$\tilde{\phi}(z, p) := p^T (\mathbf{M}(z))^{-1} p \quad (20)$$

has the following properties:

- (a) $\tilde{\phi}$ is continuous and (jointly) convex.
- (b) $\tilde{\phi}$ is nondegenerate: $\tilde{\phi}(z, p) > 0$ for all $z \in \text{int}(S)$, $p \neq 0$.
- (c) $\tilde{\phi}$ is 2-homogeneous in its second component.

Proof. Since \mathbf{M} is subject to (C0)–(C2), only convexity is not obvious. For the second directional derivative of ϕ at (z, p) in directions (ζ, π) for $\zeta \in \mathbb{R}^n$, $\pi \in \mathbb{R}^n$, we obtain,

$$D_{(z,p)}^2 \tilde{\phi}(z, p)[(\zeta, \pi), (\zeta, \pi)] = \pi^T A \pi + p^T B \pi + p^T C p, \quad (21)$$

with

$$\begin{aligned} A &:= 2\mathbf{M}(z)^{-1}, \\ B &:= -4\mathbf{M}(z)^{-1} D\mathbf{M}(z)[\zeta] \mathbf{M}(z)^{-1}, \\ C &:= 2\mathbf{M}(z)^{-1} D\mathbf{M}(z)[\zeta] \mathbf{M}(z)^{-1} D\mathbf{M}(z)[\zeta] \mathbf{M}(z)^{-1} - \mathbf{M}(z)^{-1} D^2 \mathbf{M}(z)[\zeta, \zeta] \mathbf{M}(z)^{-1}. \end{aligned}$$

We prove that the expression in (21) is nonnegative, for all admissible choices of (z, p) and (ζ, π) . Since, by condition (C1), A is symmetric positive definite, there exists a symmetric positive definite square root $A^{1/2} \in \mathbb{R}^{n \times n}$ such that $A^{1/2} A^{1/2} = A$. Further, B is symmetric. By elementary calculation, we obtain

$$\begin{aligned} D_{(z,p)}^2 \tilde{\phi}(z, p)[(\zeta, \pi), (\zeta, \pi)] &= \left| A^{1/2} \pi + \frac{1}{2} A^{-1/2} B p \right|^2 + \frac{1}{4} p^T (4C - B A^{-1} B) p \\ &= \left| A^{1/2} \pi + \frac{1}{2} A^{-1/2} B p \right|^2 - p^T \mathbf{M}(z)^{-1} D^2 \mathbf{M}(z)[\zeta, \zeta] \mathbf{M}(z)^{-1} p, \end{aligned}$$

which is nonnegative due to condition (C2). \square

Below, we need the action density to be defined up to the boundary. To this end, we replace $\tilde{\phi}$ by its lower semicontinuous envelope $\phi : S \times \mathbb{R}^n \rightarrow [0, \infty]$, defined by

$$\phi(\tilde{z}, \tilde{p}) := \liminf_{(z,p) \rightarrow (\tilde{z}, \tilde{p})} \tilde{\phi}(z, p). \quad (22)$$

Thanks to continuity of $\tilde{\phi}$, we have $\phi \equiv \tilde{\phi}$ on $\text{int}(S) \times \mathbb{R}^n$.

Example 4.2. Let $\mathbf{M}(z) = (\nabla_z^2 h_\varepsilon(z))^{-1}$ with h given in (16). Then, for $z = (z_1, z_2) \in [0, 1]^2$ and every $p = (p_1, p_2) \in \mathbb{R}^2$, we have that

$$\phi(z, p) = \begin{cases} p^T \nabla_z^2 h_\varepsilon(z) p, & \text{if } z \in (0, 1)^2, \\ \frac{p_2^2}{z_2(1-z_2)}, & \text{if } z_1 \in \{0, 1\}, z_2 \in (0, 1), p = (0, p_2), \\ \frac{p_1^2}{z_1(1-z_1)}, & \text{if } z_2 \in \{0, 1\}, z_1 \in (0, 1), p = (p_1, 0), \\ 0, & \text{if } z \in \{(0, 0), (1, 0), (1, 1), (0, 1)\} \text{ and } p = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The key step in the derivation is to observe that if z tends to a boundary point $\tilde{z} \in \partial S$ that is not a corner, then precisely one of the two eigenvalues of $\nabla_z^2 h(z)$ converges to zero, and the eigenvector for the non-vanishing eigenvalue is asymptotically parallel to ∂S at \tilde{z} .

For $\mu \in \mathcal{M}(\mathbb{R}; S)$ and $w \in \mathcal{M}(\mathbb{R}; \mathbb{R}^n)$, we define the action functional

$$\Phi(\mu, w) := \int_{\mathbb{R}} \phi(\mu(x), w(x)) \, dx. \quad (23)$$

Proposition 4.1 allows to apply Theorem 2.1 in [12] to obtain:

Proposition 4.3 (Lower semicontinuity of the action functional). *If $(\mu_k)_{k \in \mathbb{N}}$ and $(w_k)_{k \in \mathbb{N}}$ are weakly*-convergent sequences to $\mu \in \mathcal{M}(\mathbb{R}; S)$ and $w \in \mathcal{M}(\mathbb{R}; \mathbb{R}^n)$, respectively, then*

$$\liminf_{k \rightarrow \infty} \Phi(\mu_k, w_k) \geq \Phi(\mu, w).$$

4.2. Solutions to the continuity equation. Next, we investigate the structure of solutions to the (multi-component) continuity equation (19).

Since the components of μ and w are decoupled in (19), most of the results below follow from a “component-wise application” of the corresponding results in [12, 21].

Definition 4.4 (The class \mathcal{C}_T). *Given $T > 0$, define \mathcal{C}_T as the set of all curves $(\mu, w) = (\mu_t, w_t)_{t \in [0, T]}$ with the following properties:*

- (a) $(\mu_t)_{t \in [0, T]}$ is a weakly*-continuous curve in $\mathcal{M}(\mathbb{R}; S)$,
- (b) $(w_t)_{t \in [0, T]}$ is a Borel-measurable family in $\mathcal{M}(\mathbb{R}; \mathbb{R}^n)$.
- (c) For each $R > 0$,

$$\int_0^T \int_{-R}^R |w_t| dx dt < \infty,$$

- (d) (μ, w) is a distributional solution to (19) on $[0, T] \times \mathbb{R}$.

Furthermore, we denote by $\mathcal{C}_T(\hat{\mu} \rightarrow \check{\mu})$ the subset of those $(\mu, w) \in \mathcal{C}_T$ with $\mu|_{t=0} = \hat{\mu}$ and $\mu|_{t=T} = \check{\mu}$.

The continuity property (a) above imposes no restriction on the curve $(\mu_t, w_t)_{t \in [0, T]}$. Indeed, by componentwise application of Lemma 4.1 from [12], one deduces that every $(\mu_t, w_t)_{t \in [0, T]}$ satisfying (b)–(d) possesses a uniquely determined weak*-continuous representative.

Lemma 4.5 (Time rescaling). *Let $\sigma : [0, T'] \rightarrow [0, T]$ be almost everywhere equal to a diffeomorphism. Then (μ, w) is a distributional solution of (19) on $[0, T] \times \mathbb{R}$ if and only if $(\hat{\mu}, \hat{w}) := (\mu \circ \sigma, \sigma' \cdot w \circ \sigma)$ is a distributional solution of (19) on $[0, T'] \times \mathbb{R}$.*

Proof. [3, Lemma 8.1.3]. □

Lemma 4.6 (Glueing lemma). *Let $(\hat{\mu}, \hat{w}) \in \mathcal{C}_{T_1}(\mu_0 \rightarrow \mu_1)$, $(\hat{\mu}, \hat{w}) \in \mathcal{C}_{T_2}(\mu_1 \rightarrow \mu_2)$. Then the concatenation $(\mu, w) = (\mu_t, w_t)_{t \in [0, T]}$ with $T = T_1 + T_2$, defined by*

$$(\mu_t, w_t) := \begin{cases} (\hat{\mu}_t, \hat{w}_t) & \text{for } t \in [0, T_1], \\ (\hat{\mu}_{t-T_1}, \hat{w}_{t-T_1}) & \text{for } t \in (T_1, T_1 + T_2], \end{cases}$$

is an element of $\mathcal{C}_T(\mu_0 \rightarrow \mu_2)$.

Proof. This is a direct consequence of Lemma 4.5, see for instance [12]. □

Definition 4.7. The energy \mathbf{E}_T of a curve $(\mu, w) = (\mu_t, w_t)_{t \in [0, T]} \in \mathcal{C}_T$ is defined by

$$\mathbf{E}_T(\mu, w) := \int_0^T \Phi(\mu_t, w_t) dt.$$

Proposition 4.8 (Compactness in \mathcal{C}_T , part I). *Let $(\mu_k, w_k)_{k \in \mathbb{N}}$ be a sequence in \mathcal{C}_T such that for each fixed $R > 0$, the family*

$$\left\{ t \mapsto \int_{-R}^R |(w_k)_t| dx : k \in \mathbb{N} \right\}$$

of maps from $(0, T)$ into \mathbb{R}^n is k -uniformly integrable. Then, there exists a subsequence (non-relabelled) and $(\mu, w) \in \mathcal{C}_T$ such that for $k \rightarrow \infty$:

$$\begin{aligned} (\mu_k)_t &\xrightarrow{*} \mu_t \text{ weakly* in } \mathcal{M}(\mathbb{R}; S) \text{ for every } t \in [0, T], \\ w_k &\xrightarrow{*} w \text{ weakly* in } \mathcal{M}(\mathbb{R}; \mathbb{R}^n), \\ \mathbf{E}_T(\mu, w) &\leq \liminf_{k \rightarrow \infty} \mathbf{E}_T(\mu_k, w_k). \end{aligned}$$

Proof. Apply Lemma 4.5 in [12] componentwise. □

Proposition 4.9 (Compactness in \mathcal{C}_T , part II). *Let $(\mu_k, w_k)_{k \in \mathbb{N}}$ be a sequence in \mathcal{C}_T of uniformly bounded energy,*

$$\sup_{k \in \mathbb{N}} \mathbf{E}_T(\mu_k, w_k) < \infty.$$

Then the hypotheses of Proposition 4.8 are fulfilled.

Proof. To begin with, observe that thanks to continuity of \mathbf{M} by (C0), there exists a constant $C_{\mathbf{M}} > 0$ such that $\|\mathbf{M}(z)\| \leq C_{\mathbf{M}}$ for all $z \in S$. Hence $\phi(z, p) \geq C_{\mathbf{M}}^{-1}|p|^2$, for all $(z, p) \in S \times \mathbb{R}^n$. For given $R > 0$, we have that:

$$\begin{aligned} \int_0^T \left| \int_{-R}^R |(w_k)_t| dx \right|^2 dt &= \sum_{j=1}^n \int_0^T \left[\int_{-R}^R |w_{k,j}|_t dx \right]^2 dt \\ &\leq \sum_{j=1}^n \int_0^T 4R^2 \int_{\mathbb{R}} |w_{k,j}|_t^2 dx dt \leq 4R^2 C_{\mathbf{M}} \int_0^T \int_{\mathbb{R}} \phi(\mu_k, w_k) dx dt \leq 4R^2 C_{\mathbf{M}} \sup_{k \in \mathbb{N}} \mathbf{E}_T(\mu_k, w_k) \leq C < \infty. \end{aligned}$$

This proves that the family $\int_{-R}^R |(w_k)_t| dx$ is k -uniformly bounded in $L^2((0, T); \mathbb{R}^n)$. \square

4.3. Distance functional and topological properties. In this section, we prove that the distance $\mathbf{W}_{\mathbf{M}}$ with

$$\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) := [\inf \{ \mathbf{E}_1(\mu, w) : (\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1) \}]^{1/2} \quad (24)$$

is a (pseudo-) metric on $\mathcal{M}(\mathbb{R}; S)$ and investigate topological properties of $\mathbf{W}_{\mathbf{M}}$.

Proposition 4.10 (Minimizers and equivalent characterization). *The following statements hold:*

(1) *If the infimum W occurring in $\mathbf{W}_{\mathbf{M}}$ is finite, then it is attained by a curve $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$, for which one has*

$$\Phi(\mu_t, w_t) = W \quad \text{for a.e. } t \in (0, 1).$$

Consequently,

$$\mathbf{W}_{\mathbf{M}}(\mu_s, \mu_t) = |t - s| \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) \quad \forall s, t \in [0, 1].$$

(2) *There are two equivalent characterizations of $\mathbf{W}_{\mathbf{M}}$: For all $T > 0$,*

$$\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) = [\inf \{ T \mathbf{E}_T(\mu, w) : (\mu, w) \in \mathcal{C}_T(\mu_0 \rightarrow \mu_1) \}]^{1/2} \quad (25)$$

$$= \inf \left\{ \int_0^T [\Phi(\mu_t, w_t)]^{1/2} dt : (\mu, w) \in \mathcal{C}_T(\mu_0 \rightarrow \mu_1) \right\}. \quad (26)$$

Proof. The proof of (26) is essentially the same as in [12, 3], using the Rescaling Lemma 4.5. The other characterization (25) can also be obtained by this lemma using a linear rescaling of time.

For the proof of statement (1), assume that $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) = W^{1/2} < \infty$ for $W \geq 0$. Then, there exists a sequence $(\mu_k, w_k)_{k \in \mathbb{N}}$ in $\mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ with $\sup_{k \in \mathbb{N}} \mathbf{E}_1(\mu_k, w_k) < \infty$. The application of the Propositions 4.9 and 4.8 yields a limit curve $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ that is a minimizer of \mathbf{E}_1 on $\mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ due to weak*-lower semicontinuity. With (26), one deduces

$$W^{1/2} = \int_0^1 \Phi(\mu_t, w_t)^{1/2} dt,$$

and consequently, since $(0, 1) \ni t \mapsto \Phi^{1/2}(\mu_t, w_t)$ and $(0, 1) \ni t \mapsto 1$ yield equality in Hölder's inequality, $\Phi(\mu_t, w_t) = W$ for almost every $t \in (0, 1)$. \square

We are now in position to prove that $\mathbf{W}_{\mathbf{M}}$ is a distance.

Proposition 4.11 ($\mathbf{W}_{\mathbf{M}}$ is a pseudometric). *$\mathbf{W}_{\mathbf{M}}$ is a (possibly ∞ -valued) metric on the space $\mathcal{M}(\mathbb{R}; S)$.*

Proof. • **Symmetry:** This is immediate from the 2-homogeneity of ϕ and the Rescaling Lemma 4.5.

• **Definiteness:** $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) = 0$ if and only if $\mathbf{E}_1(\mu, w) = 0$ for some $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$. From positive definiteness of \mathbf{M} , this is the case if and only if $w \equiv 0$ for some $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$, hence iff $\mu_0 = \mu_1$.

- Triangle inequality: Let $\mu_0, \mu_1, \mu_2 \in \mathcal{M}(\mathbb{R}; S)$. If $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)$ or $\mathbf{W}_{\mathbf{M}}(\mu_1, \mu_2)$ is equal to $+\infty$, there is nothing to prove. If both are finite, we can use the second equivalent characterization of $\mathbf{W}_{\mathbf{M}}$ (26) and the Glueing Lemma 4.6 to obtain $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ such that $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) + \mathbf{W}_{\mathbf{M}}(\mu_1, \mu_2) = \int_0^1 \Phi(\mu_t, w_t)^{1/2} dt$. Again, invoking (26), we obtain the triangle inequality. \square

The following topological results are a consequence of the compactness results of Section 4.2, in particular of Proposition 4.9.

Proposition 4.12 (Topological properties). *The following statements hold:*

- (a) $\mathbf{W}_{\mathbf{M}}$ is lower semicontinuous in both components with respect to weak*-convergence.
- (b) Let $\mu_0 \in \mathcal{M}(\mathbb{R}; S)$ fixed, but arbitrary and let $K \subset \mathcal{M}(\mathbb{R}; S)$. If there exists $C \in \mathbb{R}$ such that $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu) \leq C$ for all $\mu \in K$, then K is relatively compact in the weak* topology.
- (c) Let $\mu_0 \in \mathcal{M}(\mathbb{R}; S)$ fixed, but arbitrary and define $\mathbf{X}[\mu_0] := \{\mu \in \mathcal{M}(\mathbb{R}; S) : \mathbf{W}_{\mathbf{M}}(\mu_0, \mu) < \infty\}$. Then, the metric space $(\mathbf{X}[\mu_0], \mathbf{W}_{\mathbf{M}})$ is complete.
- (d) $\mathbf{W}_{\mathbf{M}}^2$ is convex with respect to the linear structure of $\mathcal{M}(\mathbb{R}; S)$: If $\mu_0, \mu_1, \tilde{\mu}_0, \tilde{\mu}_1 \in \mathcal{M}(\mathbb{R}; S)$ and $\tau \in [0, 1]$, then

$$\mathbf{W}_{\mathbf{M}}^2((1 - \tau)\mu_0 + \tau\tilde{\mu}_0, (1 - \tau)\mu_1 + \tau\tilde{\mu}_1) \leq (1 - \tau)\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) + \tau\mathbf{W}_{\mathbf{M}}^2(\tilde{\mu}_0, \tilde{\mu}_1).$$

- (e) Let $\Gamma \in C^\infty(\mathbb{R})$ be nonnegative, with support in $[-1, 1]$ and $\|\Gamma\|_{L^1} = 1$, and let $\Gamma_\varepsilon(x) := \frac{1}{\varepsilon}\Gamma(\frac{x}{\varepsilon})$ for $\varepsilon > 0$. For all $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}; S)$, the following holds:

$$\begin{aligned} \mathbf{W}_{\mathbf{M}}(\mu_0 * \Gamma_\varepsilon, \mu_1 * \Gamma_\varepsilon) &\leq \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1), \\ \lim_{\varepsilon \rightarrow 0} \mathbf{W}_{\mathbf{M}}(\mu_0 * \Gamma_\varepsilon, \mu_1 * \Gamma_\varepsilon) &= \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1). \end{aligned}$$

Proof. (a) Let $(\mu_{0,k}, \mu_{1,k})_{k \in \mathbb{N}}$ be weakly* convergent to (μ_0, μ_1) as $k \rightarrow \infty$. Without loss of generality, there exists $Z \geq 0$ such that $\sup_{k \in \mathbb{N}} \mathbf{W}_{\mathbf{M}}(\mu_{0,k}, \mu_{1,k}) \leq Z$. From Proposition 4.10(1), we obtain a sequence $(\mu_k, w_k)_{k \in \mathbb{N}}$ with $(\mu_k, w_k) \in \mathcal{C}_1(\mu_{0,k} \rightarrow \mu_{1,k})$ such that $\mathbf{W}_{\mathbf{M}}^2(\mu_{0,k}, \mu_{1,k}) = \Phi((\mu_k)_t, (w_k)_t) \leq Z^2$ for almost every $t \in [0, 1]$ and all $k \in \mathbb{N}$. Hence, the requirement of Proposition 4.9 is fulfilled. The application of this proposition together with Proposition 4.8 now yields a limit curve $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ and

$$\liminf_{k \rightarrow \infty} \mathbf{W}_{\mathbf{M}}^2(\mu_{0,k}, \mu_{1,k}) = \liminf_{k \rightarrow \infty} \mathbf{E}_1(\mu_k, w_k) \geq \mathbf{E}_1(\mu, w) \geq \mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1).$$

- (b) If there exists $C \in \mathbb{R}$ such that $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu) \leq C$ for all $\mu \in K$, we can find by Proposition 4.10(1) for each $k \in \mathbb{N}$ a curve $((\mu_k)_t, (w_k)_t)_{t \in [0, 1]}$ in $\mathcal{C}_1(\mu_0 \rightarrow \mu_k)$ such that $\Phi((\mu_k)_t, (w_k)_t) \leq C^2$ for a.e. $t \in [0, 1]$ and all $k \in \mathbb{N}$. The requirement of Proposition 4.9 is again fulfilled. Its application yields in particular that $(\mu_k)_t \xrightarrow{*} \mu_t$ (on a subsequence) for all $t \in [0, 1]$ and some $(\mu_t)_{t \in [0, 1]}$.
- (c) This proof is analogous to the proof of [12, Thm. 5.7] using (a) and (b) of this proposition.
- (d) This is a consequence of convexity of the action density ϕ .
- (e) This statement can be obtained as in [12, Thm. 5.15]. \square

Proposition 4.13 (Smooth approximation of geodesics). *Assume that the mobility \mathbf{M} is induced by $h : S \rightarrow \mathbb{R}$. Let $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}; S)$ at finite distance $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) < \infty$ and such that for $i \in \{0, 1\}$:*

$$\lim_{\delta \searrow 0} \delta \int_{\mathbb{R}} [h(\mu_i) - h(\mathbf{G}_\delta * \mu_i)] dx = 0, \quad (27)$$

where \mathbf{G}_δ denotes the heat kernel

$$\mathbf{G}_s(y) := \frac{1}{\sqrt{4\pi s}} \exp\left(-\frac{y^2}{4s}\right), \quad y \in \mathbb{R}, s > 0. \quad (28)$$

Then, a geodesic curve (μ_s) connecting μ_0 and μ_1 given by Proposition 4.10 can be approximated with respect to the distance $\mathbf{W}_{\mathbf{M}}$ by the smooth curve $(\mathbf{G}_\delta * \mu_s)$, as $\delta \searrow 0$.

Proof. For each $\delta > 0$, we use the triangle inequality

$$\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) \leq \mathbf{W}_{\mathbf{M}}(\mu_0, \mathbf{G}_\delta * \mu_0) + \mathbf{W}_{\mathbf{M}}(\mathbf{G}_\delta * \mu_0, \mathbf{G}_\delta * \mu_1) + \mathbf{W}_{\mathbf{M}}(\mu_1, \mathbf{G}_\delta * \mu_1),$$

and prove that the right-hand side converges to $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)$ as $\delta \searrow 0$. Adaptation of the proof of Lemma 8.1.9 in [3] to our setting immediately yields that

$$\mathbf{W}_{\mathbf{M}}(\mathbf{G}_\delta * \mu_0, \mathbf{G}_\delta * \mu_1) \rightarrow \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1).$$

Consider now the first term above. Define for $t \in [0, 1]$

$$\begin{aligned} \tilde{\mu}_t &:= \mathbf{G}_{\delta t} * \mu_0, \\ \tilde{w}_t &:= -\delta \partial_x (\mathbf{G}_{\delta t} * \mu_0). \end{aligned}$$

Due to the smoothing property of the heat kernel, it is obvious that $(\tilde{\mu}, \tilde{w}) \in \mathcal{C}_1(\mu_0 \rightarrow \mathbf{G}_{\delta t} * \mu_0)$. For the energy of this particular curve, we obtain thanks to $\partial_x(\nabla_z h(\mathbf{G}_{\delta t} * \mu_0)) = \nabla_z^2 h(\mathbf{G}_{\delta t} * \mu_0) \partial_x(\mathbf{G}_{\delta t} * \mu_0)$ and $(\mathbf{M}(z))^{-1} = \nabla_z^2 h(z)$:

$$\begin{aligned} \mathbf{E}_1(\tilde{\mu}, \tilde{w}) &= \delta^2 \int_{[0,1] \times \mathbb{R}} \partial_x(\mathbf{G}_{\delta t} * \mu_0)^T \mathbf{M}^{-1}(\mathbf{G}_{\delta t} * \mu_0) \partial_x(\mathbf{G}_{\delta t} * \mu_0) \, d(t, x) \\ &= -\delta^2 \int_{[0,1] \times \mathbb{R}} \nabla_z h(\mathbf{G}_{\delta t} * \mu_0)^T \partial_{xx}(\mathbf{G}_{\delta t} * \mu_0) \, d(t, x) \\ &= -\delta \int_{[0,1] \times \mathbb{R}} \partial_t h(\mathbf{G}_{\delta t} * \mu_0) \, d(t, x) = \delta \int_{\mathbb{R}} [h(\mu_0) - h(\mathbf{G}_\delta * \mu_0)] \, dx \rightarrow 0, \end{aligned}$$

proving the claim. \square

Remark 4.14 (Compactly supported velocity). *With additional technical effort, one can also prove that a solution curve $(\mu, \xi) = (\mu_t, \xi_t)_{t \in [0,1]}$ to the problem*

$$\partial_t \mu + \partial_x(\mathbf{M}(\mu) \partial_x \xi) = 0, \quad \mu|_{t=0} = \mu_0, \quad \mu|_{t=1} = \mu_1,$$

can be approximated by a smooth curve $(\tilde{\mu}, \tilde{\xi})$, where $\partial_x \tilde{\xi}$ has compact support.

Under specialized conditions, an estimate of $\mathbf{W}_{\mathbf{M}}$ in terms of the second moment ℓ_2 is possible:

Proposition 4.15 (Distance and second moment). *Consider a state space $S \subset \mathbb{R}^n$ of the following form: There exists $S^\ell \in \partial S$ such that $z - S^\ell \geq 0$ (component-wise) for all $z \in S$. Assume that the mobility \mathbf{M} satisfies, in addition to (C0)–(C3), the following Lipschitz-type condition w.r.t. z :*

$$\mathbf{e}^T \mathbf{M}(z) \mathbf{e} \leq L \mathbf{e}^T (z - S^\ell), \quad \forall z \in S, \quad (29)$$

for some constant $L > 0$ and the vector $\mathbf{e} := (1, 1, \dots, 1)^T \in \mathbb{R}^n$. Then, for all $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}; S)$, one has

$$\ell_2(\mu_0 - S^\ell) \leq e^L (\ell_2(\mu_1 - S^\ell) + \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)^2).$$

Proof. Since the assertion is trivial otherwise, assume that $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) < \infty$ and $\ell_2(\mu_{1,j} - S_j^\ell) < \infty$ for all $j = 1, \dots, n$. Given $R > 0$, let $\theta_R \in C_c^\infty(\mathbb{R})$ with $\theta_R = \text{id}$ on $[-R, R]$, $\theta_R = 0$ on $\mathbb{R} \setminus [-3R, 3R]$ and $|\theta_R'(x)| \leq 1$ for all $x \in \mathbb{R}$. Observe that θ_R^2 increases to $x \mapsto x^2$ as $R \nearrow \infty$. Let $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$ be such that $\Phi(\mu_t, w_t) = \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)^2$ for almost all $t \in [0, 1]$, by Proposition 4.10. Let $s \in [0, 1]$ be arbitrary. We first obtain that

$$\int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T (\mu_s - S^\ell) \, dx - \int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T (\mu_0 - S^\ell) \, dx = - \int_0^s \int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T \partial_x w_t \, dx \, dt.$$

Using condition (C1), which yields the existence of a unique symmetric, positive definite square root $\mathbf{M}(z)^{1/2}$ of $\mathbf{M}(z)$, we get

$$\begin{aligned} - \int_0^s \int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T \partial_x w_t \, dx \, dt &= \int_0^s \int_{\mathbb{R}} 2\theta_R \theta_R' \mathbf{e}^T \mathbf{M}(\mu_t)^{1/2} \mathbf{M}(\mu_t)^{-1/2} w_t \, dx \, dt \\ &\leq \int_0^s \int_{\mathbb{R}} (\theta_R \theta_R')^2 \mathbf{e}^T \mathbf{M}(\mu_t) \mathbf{e} \, dx \, dt + \mathbf{E}_1(\mu, w), \end{aligned}$$

the last step being a consequence of the Cauchy-Schwarz and Young inequalities. Using the Lipschitz-type condition (29) and the bound on θ'_R , we end up with

$$\int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T(\mu_s - S^\ell) dx - \int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T(\mu_0 - S^\ell) dx \leq L \int_0^s \int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T(\mu_t - S^\ell) dx dt + \mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)^2.$$

Hence, by Gronwall's lemma,

$$\int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T(\mu_s - S^\ell) dx \leq e^{Ls} \left(\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) + \int_{\mathbb{R}} \theta_R^2 \mathbf{e}^T(\mu_0 - S^\ell) dx \right),$$

from which the assertion follows by monotone convergence $R \nearrow \infty$ for $s = 1$. \square

4.4. Densities at finite distance. In this section, we derive sufficient conditions under which $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)$ is finite. Throughout this section, the state space shall be a n -cuboid $S = [S^\ell, S^r]$.

Proposition 4.16 (Bounds on $\mathbf{W}_{\mathbf{M}}$ in terms of \mathbf{W}_2). *Let a mobility \mathbf{M} be given and assume that there exists a fully decoupled mobility \mathbf{M}_0 as in (14), where the scalar mobilities \mathbf{m}_j are uniformly concave, $\mathbf{m}_j'' \leq -\delta$, for some $\delta > 0$, and such that the following condition holds:*

$$\exists K > 0 : \quad A_K(z) := K\mathbf{M}_0(z)^{-1} - \mathbf{M}(z)^{-1} \in \mathbb{R}^{n \times n} \text{ is positive definite.} \quad (30)$$

Let $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}; S)$ with

$$\int_{\mathbb{R}} (\mu_0 - S^\ell) dx = m = \int_{\mathbb{R}} (\mu_1 - S^\ell) dx$$

for some $m \in [0, \infty)^n$ and $\ell_2(\mu_0 - S^\ell), \ell_2(\mu_1 - S^\ell) < \infty$. Then the following statements hold:

(a) $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1)$ is finite; in particular, one has

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \leq C[\ell_2(\mu_0 - S^\ell) + \ell_2(\mu_1 - S^\ell)] \quad (31)$$

with a constant $C > 0$ depending on m .

(b) If, moreover, for almost every $x \in \mathbb{R}$, one has $\mu_0(x), \mu_1(x) \leq \tilde{S}^r$ for $\tilde{S}^r \in \text{int}(S)$, then

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \leq \tilde{C} \sum_{j=1}^n \mathbf{W}_2^2(\mu_{0,j} - S_j^\ell, \mu_{1,j} - S_j^\ell), \quad (32)$$

with a constant $\tilde{C} > 0$ depending on m and \tilde{S}^r .

Proof. For every $(\mu, w) \in \mathcal{C}_1(\mu_0 \rightarrow \mu_1)$, one has due to condition (30) that

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \leq \mathbf{E}_1(\mu, w) \leq K \sum_{j=1}^n \int_0^1 \int_{\mathbb{R}} \frac{w_j^2}{\mathbf{m}_j(\mu_j)} dx dt.$$

Moreover, since the \mathbf{m}_j are uniformly concave, we have

$$\mathbf{m}_j(\mu_j) \geq \frac{\delta}{4}(\mu_j - S_j^\ell)(S_j^r - \mu_j) =: \tilde{\mathbf{m}}_j(\mu_j),$$

and hence

$$\mathbf{W}_{\mathbf{M}}^2(\mu_0, \mu_1) \leq K \sum_{j=1}^n \int_0^1 \int_{\mathbb{R}} \frac{w_j^2}{\tilde{\mathbf{m}}_j(\mu_j)} dx dt.$$

This estimate entails us to consider each component separately, by the same procedure as in the proof of [21, Thm. 3]. \square

In the framework of perturbations of fully decoupled mobilities (cf. Section 3.2) for $n = 2$ components, we are able to give a sufficient condition such that (30) is true.

Proposition 4.17 (Estimate on \mathbf{M}^{-1} for two components). *Assume that, for small $\varepsilon > 0$, the mobility \mathbf{M} is of the form*

$$\mathbf{M}(z) = \mathbf{M}_0(z) + \varepsilon \mathbf{M}_\varepsilon(z), \quad \text{where} \quad \mathbf{M}_0(z) := \begin{pmatrix} \mathbf{m}_1(z_1) & 0 \\ 0 & \mathbf{m}_2(z_2) \end{pmatrix},$$

with a fully decoupled mobility \mathbf{M}_0 . Assume that, in addition to (C0)–(C2), the following conditions are satisfied for some $C > 0$:

$$\begin{aligned}
(C3'a) \quad & \frac{|\mathbf{M}_{\varepsilon,11}(z)|}{\mathbf{m}_1(z_1)} < C, \\
(C3'b) \quad & \frac{|\mathbf{M}_{\varepsilon,22}(z)|}{\mathbf{m}_2(z_2)} < C, \\
(C3'c) \quad & \frac{\mathbf{m}_1(z_1)\mathbf{m}_2(z_2)}{\det \mathbf{M}(z)} < C.
\end{aligned}$$

Then, condition (30) in Proposition 4.16 holds.

Proof. We use the tr-det criterion on $A_K(z)$ and have that (omitting the argument for the sake of clarity)

$$\mathrm{tr}(A_K) > 0 \Leftrightarrow K > \frac{\mathbf{m}_1\mathbf{m}_2}{\det \mathbf{M}} \left(1 + \varepsilon \frac{\mathbf{M}_{\varepsilon,11} + \mathbf{M}_{\varepsilon,22}}{\mathbf{m}_1 + \mathbf{m}_2} \right) > 0, \quad (33)$$

$$\det(A_K) > 0 \Leftrightarrow K^2 - K \frac{2\mathbf{m}_1\mathbf{m}_2 + \varepsilon\mathbf{M}_{\varepsilon,11}(z)\mathbf{m}_2 + \varepsilon\mathbf{M}_{\varepsilon,22}\mathbf{m}_1}{\det \mathbf{M}} + \frac{\mathbf{m}_1\mathbf{m}_2}{\det \mathbf{M}} > 0. \quad (34)$$

Using the assumptions on \mathbf{M} , one easily verifies that (34) holds if

$$K > \frac{2\mathbf{m}_1\mathbf{m}_2 + \varepsilon\mathbf{M}_{\varepsilon,11}\mathbf{m}_2 + \varepsilon\mathbf{M}_{\varepsilon,22}\mathbf{m}_1}{\det \mathbf{M}} > 0. \quad (35)$$

The middle terms in (33) and (35) are strictly bounded from above by $C(1 + \varepsilon C)$, where C is the constant in (C3'a)–(C3'c). Hence, choosing $K := C(1 + \varepsilon C)$ yields the assertion. \square

5. GEODESIC CONVEXITY AND GRADIENT FLOWS

In this part of the paper, we formally establish conditions on λ -geodesic convexity of entropy functionals \mathcal{E} appearing in (1) with respect to the distance $\mathbf{W}_\mathbf{M}$. In advance of our main results, we introduce our method of proof by referring to abstract results in the literature adapted to the situation at hand.

5.1. Preliminaries. We first briefly recall the abstract setting developed in [20, 25], which is a variant of the famous “Otto calculus”. The goal is to give the metric space $(\mathcal{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M})$ a partial Riemannian structure.

A function $\mu \in \mathcal{M}(\mathbb{R}; S)$ is called *regular*, if μ is smooth and attains values in $\mathrm{int}(S)$ only. Clearly, regular functions lie dense in $\mathcal{M}(\mathbb{R}; S)$. At a regular μ , we can interpret variations $v \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ as tangent vectors to $\mathcal{M}(\mathbb{R}; S)$ at μ : each such v is associated to the curve $s \mapsto \mu + sv$ in $\mathcal{M}(\mathbb{R}; S)$. In the same spirit, each $\xi \in C^\infty(\mathbb{R}; \mathbb{R}^n)$ defines a linear functional on tangent vectors v by means of the usual pairing in $L^2(\mathbb{R}; \mathbb{R}^n)$:

$$\langle \xi, v \rangle = \int_{\mathbb{R}} \xi(x)^T v(x) \, dx.$$

Thanks to the metric structure of $(\mathcal{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M})$, there exists a distinguished injective map of linear functionals to tangent vectors at regular points μ , which is called the *Onsager operator* \mathbf{K} for the distance $\mathbf{W}_\mathbf{M}$:

$$\mathbf{K}(\mu)\xi = -\partial_x(\mathbf{M}(\mu)\partial_x\xi). \quad (36)$$

With these notions, we write (1) as an *abstract evolution equation*,

$$\partial_t \mu = -\mathbf{F}(\mu), \quad (37)$$

with the nonlinear operator $\mathbf{F} : \mathcal{M}(\mathbb{R}; S) \rightarrow \mathcal{T}\mathcal{M}(\mathbb{R}; S)$ given by

$$\mathbf{F}(\mu) := -\partial_x(\mathbf{M}(\mu)\partial_x\mathcal{E}'(\mu)) = \mathbf{K}(\mu)\mathcal{E}'(\mu). \quad (38)$$

In the framework of [20, 25], the verification of λ -geodesic convexity of \mathcal{E} with respect to the distance $\mathbf{W}_\mathbf{M}$ is based on the *Eulerian calculus* that has originally been developed in [28]; see also [11]. Theorem 5.1 below summarizes the main result of that theory.

We remark that certain hypotheses are implicitly imposed in order to justify the calculations that lead to that result. The main one is that there is a dense subset $\mathcal{M}_0 \subset \mathcal{M}(\mathbb{R}; S)$ of regular functions such that (37) possesses a smooth classical solution for each initial condition from \mathcal{M}_0 , and the associated flow maps $S^t : \mathcal{M}_0 \rightarrow \mathcal{M}(\mathbb{R}; S)$ are continuous in the topology of $(\mathcal{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M})$, for each time $t \geq 0$. It is then one of the consequences of Theorem 5.1 that $S^{(\cdot)}$ actually extends in a unique way to a continuous flow on all of

$\mathcal{M}(\mathbb{R}; S)$. Further, one needs to assume that the underlying entropy functional $\mathcal{E} : \mathcal{M}(\mathbb{R}; S) \rightarrow \mathbb{R} \cup \{\infty\}$ is proper, lower semicontinuous and bounded below.

The abstract criterion for λ -convexity is the following.

Theorem 5.1 (Condition for convexity [20, Thm. 3.6]). *Let $\lambda \in \mathbb{R}$ and let \mathcal{E} , \mathbf{F} and \mathbf{K} be defined as in (36)–(38). If*

$$\langle \xi, \mathbf{D}\mathbf{F}(\mu)\mathbf{K}(\mu)\xi \rangle - \frac{1}{2} \langle \xi, \mathbf{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle \geq \lambda \langle \xi, \mathbf{K}(\mu)\xi \rangle \quad (39)$$

holds for all regular $\mu \in \mathcal{M}(\mathbb{R}; S)$ and $\xi \in C^\infty(\mathbb{R}; \mathbb{R}^n)$ with $\partial_x \xi$ of compact support, then $\mathbf{S}^{(\cdot)}$ satisfies the evolution variational estimate (9) for \mathcal{E} and hence defines a λ -flow on $(\mathcal{M}(\mathbb{R}; S), \mathbf{W}_\mathbf{M})$. Further, \mathcal{E} is λ -geodesically convex w.r.t. $\mathbf{W}_\mathbf{M}$.

5.2. The multi-component heat equation. In this section, we apply the theory of Section 5.1 to the case of the *multi-component heat equation*,

$$\partial_t \mu = \partial_{xx} \mu, \quad (40)$$

which is (37) for $\mathbf{F}(\mu) = -\partial_{xx} \mu$. In this case, the flow maps $\mathbf{S}^t : \mathcal{M}(\mathbb{R}; S) \rightarrow \mathcal{M}(\mathbb{R}; S)$ are explicitly known:

$$\mathbf{S}^t(\mu^0) = \mathbf{G}_t * \mu^0,$$

with the heat kernel \mathbf{G} from (28), for each $t > 0$ and arbitrary initial data $\mu^0 \in \mathcal{M}(\mathbb{R}; S)$. Moreover, if μ^0 is a smooth function with values in $\text{int}(S)$ only, then it follows by classical results that $(t, x) \mapsto (\mathbf{S}^t(\mu^0))(x)$ is also smooth on $[0, \infty) \times \mathbb{R}$, and attains values in $\text{int}(S)$ only. We are thus in the framework described above and conclude the following with the help of Theorem 5.1.

Proposition 5.2 (The heat flow as a gradient flow). *Assume that $\mathbf{M} : S \rightarrow \mathbb{R}^{n \times n}$ satisfies (C0)–(C3), and that \mathbf{M} is induced by h as in (3), i.e. $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1}$ at every $z \in \text{int}(S)$, for a continuous function $h : S \rightarrow \mathbb{R}$ which is smooth on $\text{int}(S)$. Suppose that for each $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}; S)$ with $\mathbf{W}_\mathbf{M}(\mu_0, \mu_1) < \infty$, condition (27) is satisfied. Then the flow map $\mathbf{S}^{(\cdot)}$ for (40) defined above is a 0-flow on $\mathcal{M}(\mathbb{R}; S)$, and it is the gradient flow of the functional $\mathcal{H}(\mu) := \int_{\mathbb{R}} h(\mu) dx$, which is 0-geodesically convex w.r.t. $\mathbf{W}_\mathbf{M}$.*

Proof. To begin with, observe that with \mathcal{H} defined as above,

$$\mathbf{M}(\mu) \partial_x \mathcal{H}'(\mu) = \mathbf{M}(\mu) \nabla_z^2 h(\mu) \partial_x \mu = \mathbf{1} \partial_x \mu,$$

which means that (37) simplifies to (40). We verify (39) for $\lambda = 0$: for a given smooth $w : \mathbb{R} \rightarrow \mathbb{R}^n$, the relevant derivative expressions amount to

$$\begin{aligned} \mathbf{D}\mathbf{F}(\mu)[w] &= -\partial_{xx} w, \\ \mathbf{D}\mathbf{K}(\mu)[w]\xi &= -\partial_x (\mathbf{D}\mathbf{M}(\mu)[w] \partial_x \xi). \end{aligned} \quad (41)$$

We substitute this into the left-hand side of (39) and integrate by parts to obtain

$$\begin{aligned} \langle \xi, \mathbf{D}\mathbf{F}(\mu)[\mathbf{K}(\mu)\xi] \rangle - \frac{1}{2} \langle \xi, \mathbf{D}\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle &= \langle \xi, \partial_{xxx} (\mathbf{M}(\mu) \partial_x \xi) \rangle - \frac{1}{2} \langle \xi, \partial_x (\mathbf{D}\mathbf{M}(\mu)[\partial_{xx} \mu] \partial_x \xi) \rangle \\ &= -\frac{1}{2} \langle \partial_x \xi, \mathbf{D}^2 \mathbf{M}(\mu)[\partial_x \mu, \partial_x \mu] \partial_x \xi \rangle + \langle \partial_{xx} \xi, \mathbf{M}(\mu) \partial_{xx} \xi \rangle, \end{aligned}$$

which is nonnegative because of (C1) and (C2). \square

5.3. Internal energy functionals. We now study geodesic convexity of more general functionals of the form

$$\mathcal{E}(\mu) = \int_{\mathbb{R}} f(\mu(x)) dx, \quad (42)$$

with a smooth function $f : \text{int}(S) \rightarrow \mathbb{R}$. For brevity, we call these functionals *internal energies*, regardless of their actual interpretation in physics or other sciences. Our main result is Proposition 5.3 below, which is a further generalization of the generalized McCann condition established by Carrillo *et al.* [8] for *scalar* nonlinear mobilities ($n = 1$).

5.3.1. *A generalized McCann condition.* The main result of this section is the following sufficient criterion for 0-contractivity of the flow generated by the evolution equation

$$\partial_t \mu = \partial_x (\mathbf{L}(\mu) \partial_x \mu), \quad \text{with} \quad \mathbf{L}(z) = \mathbf{M}(z) \nabla_z^2 f(z), \quad (43)$$

which is (1) for \mathcal{E} from (42), i.e., the formal gradient flow of \mathcal{E} in $\mathbf{W}_{\mathbf{M}}$.

Proposition 5.3 (Multi-component McCann condition). *Given a mobility matrix \mathbf{M} that satisfies (C0)–(C2) and a functional \mathcal{E} of the form (42), assume that for all $z \in \text{int}(S)$ and all $v, \zeta, \beta \in \mathbb{R}^n$ (omitting the argument z from $\mathbf{M} = \mathbf{M}(z)$ and from $\mathbf{L} = \mathbf{M}(z) \nabla_z^2 f(z)$):*

$$\begin{aligned} 0 \leq & -\frac{1}{2} v^T \mathbf{D}^2 \mathbf{M}[\zeta, \mathbf{L}\zeta] v + \beta^T \mathbf{L} \mathbf{M} \beta \\ & + \beta^T (\mathbf{L} \mathbf{D} \mathbf{M}[\zeta] - \mathbf{D} \mathbf{M}[\mathbf{L}\zeta]) v + v^T \mathbf{D} \mathbf{L}[\zeta] (\mathbf{D} \mathbf{M}[\zeta] v + \mathbf{M} \beta) - v^T \mathbf{D} \mathbf{L}[\mathbf{D} \mathbf{M}[\zeta] v + \mathbf{M} \beta] \zeta. \end{aligned} \quad (44)$$

Then, under the assumption of sufficient regularity of the associated flow generated by (43), the functional \mathcal{E} is 0-geodesically convex w.r.t. the distance $\mathbf{W}_{\mathbf{M}}$.

Proof. This is another application of Theorem 5.1. Let therefore $\mu \in \mathcal{M}(\mathbb{R}; S)$ be regular and $\xi, w \in C^\infty(\mathbb{R}; \mathbb{R}^n)$, $\partial_x \xi$ with compact support. Observe that

$$\begin{aligned} \mathbf{F}(\mu) &= -\partial_x (\mathbf{L}(\mu) \mu_x), \\ \mathbf{D} \mathbf{F}(\mu)[w] &= -\partial_x (\mathbf{D} \mathbf{L}(\mu)[w] \mu_x) - \partial_x (\mathbf{L}(\mu) w_x), \end{aligned}$$

and in addition, (41) holds. Hence, integrating by parts, we obtain

$$\begin{aligned} & \langle \xi, \mathbf{D} \mathbf{F}(\mu) \mathbf{K}(\mu) \xi \rangle - \frac{1}{2} \langle \xi, \mathbf{D} \mathbf{K}(\mu) [\mathbf{F}(\mu)] \xi \rangle \\ &= -\langle \xi_x, \mathbf{D} \mathbf{L}(\mu) [\partial_x (\mathbf{M}(\mu) \xi_x)] \mu_x \rangle + \langle \xi_{xx}, \mathbf{L}(\mu) \partial_x (\mathbf{M}(\mu) \xi_x) \rangle + \langle \xi_{xx}, \mathbf{D} \mathbf{L}(\mu) [\mu_x] \partial_x (\mathbf{M}(\mu) \xi_x) \rangle \\ & - \frac{1}{2} \langle \xi_{xx}, \mathbf{D} \mathbf{M}(\mu) [\mathbf{L}(\mu) \mu_x] \xi_x \rangle - \frac{1}{2} \langle \xi_x, \mathbf{D} \mathbf{M}(\mu) [\mathbf{L}(\mu) \mu_x] \xi_{xx} \rangle - \frac{1}{2} \langle \xi_x, \mathbf{D}^2 \mathbf{M}(\mu) [\mu_x, \mathbf{L}(\mu) \mu_x] \xi_x \rangle \\ &= -\frac{1}{2} \langle \xi_x, \mathbf{D}^2 \mathbf{M}(\mu) [\mu_x, \mathbf{L}(\mu) \mu_x] \xi_x \rangle - \langle \xi_{xx}, \mathbf{D} \mathbf{M}(\mu) [\mathbf{L}(\mu) \mu_x] \xi_x \rangle - \langle \xi_x, \mathbf{D} \mathbf{L}(\mu) [\mathbf{D} \mathbf{M}(\mu) [\mu_x] \xi_x + \mathbf{M}(\mu) \xi_{xx}] \mu_x \rangle \\ & + \langle \xi_{xx}, \mathbf{L}(\mu) (\mathbf{D} \mathbf{M}(\mu) [\mu_x] \xi_x + \mathbf{M}(\mu) \xi_{xx}) \rangle + \langle \xi_{xx}, \mathbf{D} \mathbf{L}(\mu) [\mu_x] (\mathbf{D} \mathbf{M}(\mu) [\mu_x] \xi_x + \mathbf{M}(\mu) \xi_{xx}) \rangle. \end{aligned}$$

Condition (44) now implies pointwise nonnegativity (substitute $v := \xi_x(x)$, $\beta := \xi_{xx}(x)$, $\zeta := \mu_x(x)$ for $x \in \mathbb{R}$) and consequently (39) for $\lambda = 0$. \square

Remark 5.4 (Diagonal mobility). *In the case of a fully decoupled mobility matrix*

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & 0 \\ 0 & \mathbf{m}_2(z_2) \end{pmatrix}$$

for $n = 2$ components, where in general $\frac{1}{2} \mathbf{m}_j'' \mathbf{m}_j + (\mathbf{m}_j')^2 \neq 0$, the generalized McCann condition (44) is equivalent to

$$\partial_{11} f(z) \geq 0, \quad \partial_{22} f(z) \geq 0, \quad \partial_{12} f(z) = 0,$$

since (44) reads in this case

$$\begin{aligned} 0 \geq & \left[\frac{1}{2} v_1^2 \zeta_1^2 \mathbf{m}_1'' \mathbf{m}_1 - \beta_1^2 \mathbf{m}_1^2 \right] \partial_{11} f + \left[\frac{1}{2} v_2^2 \zeta_2^2 \mathbf{m}_2'' \mathbf{m}_2 - \beta_2^2 \mathbf{m}_2^2 \right] \partial_{22} f \\ & + \left[\frac{1}{2} v_1^2 \zeta_1 \zeta_2 \mathbf{m}_1'' \mathbf{m}_1 + \frac{1}{2} v_2^2 \zeta_1 \zeta_2 \mathbf{m}_2'' \mathbf{m}_2 - 2\beta_1 \beta_2 \mathbf{m}_1 \mathbf{m}_2 - 2v_1 v_2 \zeta_1 \zeta_2 \mathbf{m}_1' \mathbf{m}_2' \right] \partial_{12} f \\ & + [v_1^2 (\mathbf{m}_1')^2 + v_2^2 (\mathbf{m}_2')^2] \zeta_1 \zeta_2 + 2v_1 \beta_1 \zeta_2 \mathbf{m}_1' \mathbf{m}_1 + 2v_2 \beta_2 \zeta_1 \mathbf{m}_2' \mathbf{m}_2 - 2v_2 \beta_1 \zeta_2 \mathbf{m}_2' \mathbf{m}_1 - 2v_1 \beta_2 \zeta_1 \mathbf{m}_1' \mathbf{m}_2 \partial_{12} f. \end{aligned}$$

Imposing e.g. $\beta = 0$, $v_1 = 1$, $v_2 = 0$ and $\zeta_1 = 1$, one obtains

$$0 \geq \frac{1}{2} \mathbf{m}_1'' \mathbf{m}_1 \partial_{11} f + \left[\frac{1}{2} \mathbf{m}_1'' \mathbf{m}_1 + (\mathbf{m}_1')^2 \right] \zeta_2 \partial_{12} f,$$

from which necessarily $\partial_{12} f(\mu) = 0$ follows. Hence, the only possible choice is $f(z) := \psi_1(z_1) + \psi_2(z_2)$ with convex functions ψ_1, ψ_2 . We solely recover the generalized McCann condition for $n = 1$ (cf. [8]) for each of the two components separately if \mathbf{M} is fully decoupled.

5.3.2. Perturbation results and examples. This paragraph is devoted to examples satisfying condition (44) of Proposition 5.3. In particular, we investigate suitable perturbations of the entropies having the heat flow as gradient flow, cf. Proposition 5.2. We first start with a more general result involving perturbations of compact support in $\text{int}(S)$ and continue with a specific example where the support of the perturbation extends to all of S .

Proposition 5.5 (Perturbations of compact support). *Let a mobility \mathbf{M} satisfy the conditions (C0)–(C3) and the stronger condition (C2') and be induced by h as in (3). For $\alpha, \tilde{\varepsilon} > 0$ and $g \in C_c^\infty(\text{int}(S))$, define $f(z) := \alpha h(z) + \tilde{\varepsilon} g(z)$ and \mathcal{E} according to (42). Then, for $\tilde{\varepsilon} > 0$ sufficiently small, the generalized McCann condition (44) is satisfied.*

Proof. If $z \notin \text{supp } g$, the conditions (C1) and (C2') directly yield the claim. Furthermore, there exists a constant $\delta_g > 0$ such that for all $z \in \text{supp } g$, one has

$$\begin{aligned} \beta^T \mathbf{D}^2 \mathbf{M}(z) [\zeta, \zeta] \beta &\leq -\delta_g |\beta|^2 |\zeta|^2, \\ -\gamma^T \mathbf{M}(z) \gamma &\leq -\delta_g |\gamma|^2, \end{aligned}$$

for all $\beta, \gamma, \zeta \in \mathbb{R}^n$. Hence, by continuity, we obtain for the r.h.s. in (44), recalling

$$\mathbf{L}(z) = \mathbf{M}(z) \nabla_z^2 f(z) = \alpha \mathbf{1} + \tilde{\varepsilon} \nabla_z^2 g(z) :$$

$$\begin{aligned} & -\frac{1}{2} v^T \mathbf{D}^2 \mathbf{M}[\zeta, \mathbf{L}\zeta] v + \beta^T \mathbf{L} \mathbf{M} \beta \\ & + \beta^T (\mathbf{L} \mathbf{D} \mathbf{M}[\zeta] - \mathbf{D} \mathbf{M}[\mathbf{L}\zeta]) v + v^T \mathbf{D} \mathbf{L}[\zeta] (\mathbf{D} \mathbf{M}[\zeta] v + \mathbf{M} \beta) - v^T \mathbf{D} \mathbf{L} [\mathbf{D} \mathbf{M}[\zeta] v + \mathbf{M} \beta] \zeta \\ & \geq \frac{\alpha}{2} \delta_g |\zeta|^2 |v|^2 + \alpha \delta_g |\beta|^2 - \tilde{\varepsilon} C_{g, \mathbf{M}} (|\zeta|^2 |v|^2 + |\beta|^2 + |\zeta| |v| |\beta|), \end{aligned}$$

with a constant $C_{g, \mathbf{M}} > 0$. Using Young's inequality, one immediately deduces that the r.h.s. is nonnegative and thus (44) is satisfied, provided that $\tilde{\varepsilon} \leq \frac{\alpha \delta_g}{3 C_{g, \mathbf{M}}}$. \square

We conclude this section with a specific example such that the support of the perturbation g extends to all of S .

Example 5.6 (Non-compactly supported perturbations). *Let \mathbf{M} be induced by h from (15) & (16):*

$$\begin{aligned} h(z) &:= z_1 \log(z_1) + (1 - z_1) \log(1 - z_1) + z_2 \log(z_2) + (1 - z_2) \log(1 - z_2) + \varepsilon d_1 d_2, \\ d_j &:= z_j (1 - z_j), \end{aligned}$$

and $\varepsilon > 0$ chosen so small such that the conditions (C0)–(C3) and (C2') are satisfied. Define furthermore $\tilde{g} : [0, \frac{1}{4}]^2 \rightarrow \mathbb{R}$ by

$$\tilde{g}(m_1, m_2) := \exp \left(-\frac{1}{m_1} - \frac{1}{m_2} \right)$$

for all $0 < m_1, m_2 \leq \frac{1}{4}$, and $\tilde{g}(m_1, 0) = 0 = \tilde{g}(0, m_2)$. Consider now for $\tilde{\varepsilon} > 0$ the map $f(z) := h(z) + \tilde{\varepsilon} \tilde{g}(d_1, d_2)$ and the functional \mathcal{E} according to (42). Then, for $\tilde{\varepsilon} > 0$ sufficiently small, the generalized McCann condition (44) is satisfied.

Our idea of proof relies on the structure of \mathbf{M} in this particular case (cf. Section 3.2): There exists a positive rational function $r_1 : (0, \frac{1}{4}]^2 \rightarrow (0, \infty)$ with $\lim_{\tilde{m} \rightarrow 0} r_1(m_1, \tilde{m}) = 0 = \lim_{\tilde{m} \rightarrow 0} r_1(\tilde{m}, m_2)$ for all $(m_1, m_2) \in (0, \frac{1}{4}]^2$, such that

$$\frac{1}{2} \beta^T \mathbf{D}^2 \mathbf{M}(z) [\zeta, \zeta] \beta - \gamma^T \mathbf{M}(z) \gamma \leq -r_1(d_1, d_2) (|\zeta|^2 |\beta|^2 + |\gamma|^2).$$

Furthermore, there exists another rational function $r_2 : (0, \frac{1}{4}]^2 \rightarrow [0, \infty)$ such that the following estimate on the r.h.s. in condition (44) is possible:

$$\begin{aligned} & -\frac{1}{2} v^T \mathbf{D}^2 \mathbf{M}[\zeta, \mathbf{L}\zeta] v + \beta^T \mathbf{L} \mathbf{M} \beta \\ & + \beta^T (\mathbf{L} \mathbf{D} \mathbf{M}[\zeta] - \mathbf{D} \mathbf{M}[\mathbf{L}\zeta]) v + v^T \mathbf{D} \mathbf{L}[\zeta] (\mathbf{D} \mathbf{M}[\zeta] v + \mathbf{M} \beta) - v^T \mathbf{D} \mathbf{L} [\mathbf{D} \mathbf{M}[\zeta] v + \mathbf{M} \beta] \zeta \\ & \geq (r_1(d_1, d_2) - \tilde{g}(d_1, d_2) r_2(d_1, d_2)) (|\zeta|^2 |\beta|^2 + |\gamma|^2). \end{aligned}$$

Since for all $(m_1, m_2) \in (0, \frac{1}{4}]^2$, one has

$$\lim_{\tilde{m} \rightarrow 0} \tilde{g}(\tilde{m}, m_2) \frac{r_2(\tilde{m}, m_2)}{r_1(\tilde{m}, m_2)} = 0 = \lim_{\tilde{m} \rightarrow 0} \tilde{g}(m_1, \tilde{m}) \frac{r_2(m_1, \tilde{m})}{r_1(m_1, \tilde{m})},$$

we find $\tilde{\varepsilon}_0 > 0$ sufficiently small such that (44) holds for all $0 < \tilde{\varepsilon} \leq \tilde{\varepsilon}_0$.

5.4. The potential energy. In this section, we study λ -convexity of the regularized *potential energy* functional

$$\mathcal{V}(\mu) = \int_{\mathbb{R}} [\alpha h(\mu) + \rho(x)^T \mu] \, dx, \quad (45)$$

which has a density depending explicitly on the spatial variable x . Here, $\mathbf{W}_{\mathbf{M}}$ and h are as in Proposition 5.2 and $\alpha > 0$, $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ are fixed. The flow associated to \mathcal{V} is generated by the following (regularized) nonlinear *transport equation*:

$$\partial_t \mu = \alpha \partial_{xx} \mu + \partial_x (\mathbf{M}(\mu) \partial_x \rho). \quad (46)$$

5.4.1. Convexity. A sufficient condition on convexity of those entropies is the following:

Proposition 5.7 (Convexity for the regularized potential energy functional). *Let \mathcal{V} be of the form (45) with h , α and ρ as mentioned above, let $\mathbf{M} = (\nabla_z^2 h)^{-1}$ be as in Proposition 5.2 and $\lambda \in \mathbb{R}$ be fixed. If for all $z \in \text{int}(S)$ and all $v, \zeta \in \mathbb{R}^n$, $q^1, q^2 \in \mathbb{B}_R(0)$, $R := \|\rho\|_{C^2}$, the condition*

$$\begin{aligned} 0 \leq & -\frac{\alpha}{2} v^T D^2 \mathbf{M}[\zeta, \zeta] v - \lambda v^T \mathbf{M} v \\ & - \frac{1}{2} v^T D^2 \mathbf{M}[\zeta, \mathbf{M} q^1] v + v^T D^2 \mathbf{M}[\zeta, \mathbf{M} v] q^1 + v^T D \mathbf{M}[\mathbf{M} q^2] v \end{aligned} \quad (47)$$

is satisfied, then \mathcal{V} is λ -geodesically convex w.r.t. the distance $\mathbf{W}_{\mathbf{M}}$ under the assumption of sufficient regularity of the associated flow generated by (46).

Proof. The method of proof is similar to that of Proposition 5.3. Here, one gets

$$\begin{aligned} \mathbf{F}(\mu) &= -\alpha \partial_{xx} \mu - \partial_x (\mathbf{M}(\mu) \rho_x), \\ D\mathbf{F}(\mu)[w] &= -\alpha \partial_{xx} w - \partial_x (D\mathbf{M}(\mu)[w] \rho_x). \end{aligned}$$

Consequently, performing essentially the same calculations as in the proofs of the Propositions 5.2 and 5.3,

$$\begin{aligned} & -\frac{1}{2} \langle \xi, D\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle + \langle \xi, D\mathbf{F}(\mu)\mathbf{K}(\mu)\xi \rangle - \lambda \langle \xi, \mathbf{K}(\mu)\xi \rangle \\ &= -\frac{\alpha}{2} \langle \xi_x, D^2 \mathbf{M}(\mu)[\mu_x, \mu_x]\xi_x \rangle - \frac{1}{2} \langle \xi_x, D^2 \mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\rho_x]\xi_x \rangle \\ &+ \alpha \langle \xi_{xx}, \mathbf{M}(\mu)\xi_{xx} \rangle + \langle \xi_x, D^2 \mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\xi_x]\rho_x \rangle + \langle \xi_x, D\mathbf{M}(\mu)[\mathbf{M}(\mu)\xi_x]\rho_{xx} \rangle \\ &- \lambda \langle \xi_x, \mathbf{M}(\mu)\xi_x \rangle. \end{aligned}$$

We use the fact that for all $\gamma, q^1, v \in \mathbb{R}^n$ and all $z \in \text{int}(S)$, one has due to symmetry of the third-order tensor $D^3 h$:

$$\gamma^T D\mathbf{M}(z)[\mathbf{M}(\mu)q^1]v = -D^3 h(z)[\mathbf{M}(z)\gamma, \mathbf{M}(z)q^1, \mathbf{M}(z)v] = \gamma^T D\mathbf{M}(z)[\mathbf{M}(z)v]q^1.$$

Hence, we obtain

$$\begin{aligned} & \langle \xi, D\mathbf{F}(\mu)\mathbf{K}(\mu)\xi \rangle - \frac{1}{2} \langle \xi, D\mathbf{K}(\mu)[\mathbf{F}(\mu)]\xi \rangle - \lambda \langle \xi, \mathbf{K}(\mu)\xi \rangle \\ &= -\frac{\alpha}{2} \langle \xi_x, D^2 \mathbf{M}(\mu)[\mu_x, \mu_x]\xi_x \rangle + \alpha \langle \xi_{xx}, \mathbf{M}(\mu)\xi_{xx} \rangle - \lambda \langle \xi_x, \mathbf{M}(\mu)\xi_x \rangle \\ &- \frac{1}{2} \langle \xi_x, D^2 \mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\rho_x]\xi_x \rangle + \langle \xi_x, D^2 \mathbf{M}(\mu)[\mu_x, \mathbf{M}(\mu)\xi_x]\rho_x \rangle + \langle \xi_x, D\mathbf{M}(\mu)[\mathbf{M}(\mu)\rho_{xx}]\xi_x \rangle, \end{aligned}$$

which is nonnegative due to condition (47) and (C1) (substitute $v := \xi_x(x)$, $\zeta := \mu_x(x)$, $q^1 := \rho_x(x)$, $q^2 := \rho_{xx}(x)$ for $x \in \mathbb{R}$) and hence implies (39). \square

5.4.2. *The case of fully decoupled mobility.* In this paragraph, we consider the case of a fully decoupled mobility (cf. Section 3.1)

$$\mathbf{M}(z) = \begin{pmatrix} \mathbf{m}_1(z_1) & & \\ & \ddots & \\ & & \mathbf{m}_n(z_n) \end{pmatrix},$$

on the n -cuboid $S = [S^\ell, S^r]$. We shall assume that the scalar mobilities \mathbf{m}_j are such that

- $\mathbf{m}_j \in C^2([S_j^\ell, S_j^r])$,
- $\mathbf{m}_j(s) > 0$ for $s \in (S_j^\ell, S_j^r)$ and $\mathbf{m}_j(S_j^\ell) = \mathbf{m}_j(S_j^r) = 0$,
- $\mathbf{m}_j''(s) \leq 0$ for $s \in [S_j^\ell, S_j^r]$.

Recall that \mathbf{M} is of the special form (3) $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1}$, where

$$h(z) = \sum_{j=1}^n h_j(z_j),$$

h_j being a second primitive of $\frac{1}{\mathbf{m}_j}$.

Proposition 5.8 (λ -convexity of the potential energy). *For a fully decoupled mobility \mathbf{M} as mentioned above, fix $\alpha > 0$ and $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and consider the regularized potential energy functional \mathcal{V} defined in (45).*

(a) *Let $\bar{z} \in S$ and $\mu^0 \in \mathcal{M}(\mathbb{R}; S)$ such that $\mu^0 - \bar{z} \in H^1(\mathbb{R}; \mathbb{R}^n)$ and such that μ^0 attains values in $\text{int}(S)$ only. Then, the initial-value problem for (46)*

$$\partial_t \mu = \alpha \partial_{xx} \mu + \partial_x (\mathbf{M}(\mu) \partial_x \rho), \quad \mu(0, \cdot) = \mu^0, \quad (48)$$

possesses a unique local-in-time classical solution $\mu : [0, T] \rightarrow \mathcal{M}(\mathbb{R}; S)$ with $\mu - \bar{z} \in C^0([0, T]; H^1(\mathbb{R}; \mathbb{R}^n))$, where $T = T(\mu^0, \rho) > 0$.

(b) *There exists $C = C(\rho) > 0$ such that condition (47) in Proposition 5.7 is satisfied for all $\lambda \leq -C(\frac{1}{\alpha} + 1)$. Hence, Proposition 5.7 is applicable and yields λ -convexity of the potential energy \mathcal{V} .*

Proof. (a) See Appendix A.

(b) We proceed similarly to [22] and observe that for all $z \in \text{int}(S)$ and all $v, \zeta \in \mathbb{R}^n$, $q^1, q^2 \in \overline{\mathbb{B}_R(0)}$, $R := \|\rho\|_{C^2}$, one has

$$\begin{aligned} & -\frac{\alpha}{2} v^T D^2 \mathbf{M}[\zeta, \zeta] v - \frac{1}{2} v^T D^2 \mathbf{M}[\zeta, \mathbf{M} q^1] v + v^T D^2 \mathbf{M}[\zeta, \mathbf{M} v] q^1 + v^T D \mathbf{M}[\mathbf{M} q^2] v \\ &= \sum_{j=1}^n \left[-\frac{\alpha}{2} \mathbf{m}_j''(z_j) \zeta_j^2 + \frac{1}{2} \mathbf{m}_j''(z_j) \mathbf{m}_j(z_j) q_j^1 \zeta_j + \mathbf{m}_j'(z_j) \mathbf{m}_j(z_j) q_j^2 \right] v_j^2 \\ &\geq \sum_{j=1}^n \left[-\frac{1}{8\alpha} |\mathbf{m}_j''(z_j)| \mathbf{m}_j(z_j)^2 (q_j^1)^2 + \mathbf{m}_j'(z_j) \mathbf{m}_j(z_j) q_j^2 \right] v_j^2, \end{aligned}$$

the last step being a consequence of Young's inequality. Using the bounds on \mathbf{m}_j , q^1 and q^2 , we obtain

$$\sum_{j=1}^n \left[-\frac{1}{8\alpha} |\mathbf{m}_j''(z_j)| \mathbf{m}_j(z_j) (q_j^1)^2 + \mathbf{m}_j'(z_j) q_j^2 \right] \mathbf{m}_j(z_j) v_j^2 \geq -\sum_{j=1}^n \|\mathbf{m}_j\|_{C^2} R \left[\frac{\|\mathbf{m}_j\|_{C^2} R}{8\alpha} + 1 \right] \mathbf{m}_j(z_j) v_j^2.$$

Obviously, for all $\lambda \leq -\max_j \|\mathbf{m}_j\|_{C^2} R \left[\frac{\|\mathbf{m}_j\|_{C^2} R}{8\alpha} + 1 \right]$, (47) holds. □

6. EXISTENCE OF WEAK SOLUTIONS

In this section, we prove the existence of weak solutions for a class of initial-value problems of the form (1). More specifically, we consider the case of a fully decoupled mobility \mathbf{M} but allow for coupling inside the driving entropy \mathcal{E} . Note that, by Remark 5.4, the functional \mathcal{E} will in general *not* be geodesically convex.

6.1. Setting and basic properties. We again consider as state space a n -cuboid $S = [S^\ell, S^r] \subset \mathbb{R}^n$ and let $h : S \rightarrow \mathbb{R}$, $h(z) = \sum_{j=1}^n h_j(z_j)$, where for all $j = 1, \dots, n$:

(H0) h_j is α -Hölder continuous on $[S_j^\ell, S_j^r]$ for some $\alpha \in (0, 1]$ and smooth on (S_j^ℓ, S_j^r) ,

(H1) h_j is strictly convex,

(H2) $\lim_{s \searrow S_j^\ell} h_j''(s) = +\infty = \lim_{s \nearrow S_j^r} h_j''(s)$.

(H3) $\frac{1}{h_j''}$ is concave and can be extended at the boundary $\{S_j^\ell, S_j^r\}$ to a function in $C^2([S_j^\ell, S_j^r])$.

Obviously, the induced fully decoupled mobility \mathbf{M} as in Section 5.4.2 satisfies the requirements of that section, in particular also (C0)–(C3), if h satisfies (H0)–(H3).

Furthermore, let $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and $f : S \rightarrow \mathbb{R}$ such that

(F) f is smooth and uniformly convex, i.e. $\nabla_z^2 f(z) \geq C_f \mathbb{1}$ for all $z \in S$ and some $C_f > 0$.

We introduce a *reference state* $\bar{z} \in S$, i.e. a constant level relatively to which certain quantities (e.g. the mass of an element in $\mathcal{M}(\mathbb{R}; S)$) will be measured. We distinguish two qualitatively different cases:

(A) Reference state $\bar{z} = S^\ell$.

(B) Reference state $\bar{z} \in \text{int}(S)$.

The respective case will be indicated with (A) and/or (B) in definitions and statements. Note that in case (A), the function $\mu - \bar{z}$ is nonnegative for each $\mu \in \mathcal{M}(\mathbb{R}; S)$.

Definition 6.1 (Heat and driving entropy). *Let \bar{z}, f, h, η be as mentioned above. In case (A), let $\mu^0 \in \mathcal{M}(\mathbb{R}; S)$ be such that $m := \|\mu^0 - \bar{z}\|_{L^1} \in (0, \infty)$. Define the heat entropy functional by*

$$\mathcal{H}(\mu) = \int_{\mathbb{R}} h_{\bar{z}}(\mu) \, dx,$$

where

(A) $h_{\bar{z}} := h(z) - h(\bar{z})$,

(B) $h_{\bar{z}}(z) := h(z) - h(\bar{z}) - (z - \bar{z})^T \nabla_z h(\bar{z})$.

The driving entropy functional $\mathcal{E} : \mathcal{M}(\mathbb{R}; S) \rightarrow \mathbb{R} \cup \{\infty\}$ is defined by

$$\mathcal{E}(\mu) = \begin{cases} \int_{\mathbb{R}} [f(\mu) - f(\bar{z}) - (\mu - \bar{z})^T \nabla_z f(\bar{z}) + \mu^T \eta] \, dx, & \text{if } \mu \in \mathbf{X}_{\bar{z}}, \\ +\infty, & \text{otherwise,} \end{cases}$$

where

(A) $\mathbf{X}_{\bar{z}} := \{\mu \in \mathcal{M}(\mathbb{R}; S) : \|\mu - \bar{z}\|_{L^1} = m, \ell_2(\mu - \bar{z}) < \infty\}$,

(B) $\mathbf{X}_{\bar{z}} := \{\mu \in \mathcal{M}(\mathbb{R}; S) : \|\mu - \bar{z}\|_{L^2} < \infty\}$.

Note that in both cases, $h_{\bar{z}}(\bar{z}) = 0$ and $h_{\bar{z}}$ is strictly convex with $\nabla_z^2 h_{\bar{z}} = \nabla_z^2 h$. In case (B), $h_{\bar{z}}$ is nonnegative.

Example 6.2. (a) The paradigmatic example for h satisfying (H0)–(H3) is given by

$$h_j(s) = \begin{cases} (s - S_j^\ell) \log(s - S_j^\ell) + (S_j^r - s) \log(S_j^r - s) - (S_j^r - S_j^\ell) \log(S_j^r - S_j^\ell), & \text{if } s \in (S_j^\ell, S_j^r), \\ 0, & \text{if } s \in \{S_j^\ell, S_j^r\}, \end{cases}$$

yielding

$$\mathbf{m}_j(s) = \frac{1}{S_j^r - S_j^\ell} (s - S_j^\ell)(S_j^r - s).$$

(b) An admissible choice for f is

$$f(z) = \frac{1}{2} z^T Q z + \varepsilon r(z),$$

where $Q \in \mathbb{R}^{n \times n}$ is symmetric positive definite, $r : S \rightarrow \mathbb{R}$ is smooth and $\varepsilon \geq 0$ is such that $Q + \varepsilon \nabla_z^2 r(z)$ is positive definite for all $z \in S$.

Proposition 6.3 (Properties of heat and driving entropy (A)+(B)). *The following statements hold:*

(a) \mathcal{H} is finite on $\mathbf{X}_{\bar{z}}$.

- (b) For all $\mu_0, \mu_1 \in \mathbf{X}_{\bar{z}}$ with $\mathbf{W}_{\mathbf{M}}(\mu_0, \mu_1) < \infty$, condition (27) holds for $h_{\bar{z}}$ in place of h .
(c) The Lipschitz-type condition (29) holds.
(d) There exist constants $\underline{C}, \bar{C} > 0$ such that for all $\mu \in \mathbf{X}_{\bar{z}}$, the following holds:

$$\underline{C}(\|\mu - \bar{z}\|_{L^2}^2 - 1) \leq \mathcal{E}(\mu) \leq \bar{C}(\|\mu - \bar{z}\|_{L^2}^2 + 1).$$

In particular, \mathcal{E} is finite on $\mathbf{X}_{\bar{z}}$.

- (e) If $\mu_k - \bar{z} \rightharpoonup \mu - \bar{z}$ weakly in $L^2(\mathbb{R}; \mathbb{R}^n)$, then

$$\mathcal{E}(\mu) dx \leq \liminf_{k \rightarrow \infty} \mathcal{E}(\mu_k) dx.$$

Proof. (a) We distinguish both cases.

- (A) Due to α -Hölder continuity of h , there exists $C > 0$ such that for all $z \in S$:

$$|h_{\bar{z}}(z)| \leq C \sum_{j=1}^n |z_j - \bar{z}_j|^\alpha.$$

By Hölder's inequality, we then deduce for $\mu \in \mathbf{X}_{\bar{z}}$:

$$|\mathcal{H}(\mu)| \leq C \sum_{j=1}^n \int_{\mathbb{R}} |\mu_j - \bar{z}_j|^\alpha dx \leq C \sum_{j=1}^n \left(\int_{\mathbb{R}} (\mu_j - \bar{z}_j)(x^2 + 1) dx \right)^\alpha \left(\int_{\mathbb{R}} (x^2 + 1)^{\frac{\alpha}{\alpha-1}} dx \right)^{1-\alpha},$$

which is finite thanks to the definition of $\mathbf{X}_{\bar{z}}$.

- (B) Obviously, since h is smooth in a neighbourhood of \bar{z} and bounded on the whole of S , there exists $C > 0$ such that for all $z \in S$:

$$h_{\bar{z}}(z) \leq C|z - \bar{z}|^2,$$

which proves the claim.

- (b) (A) Thanks to the properties of the heat kernel, $\mathbf{G}_\delta * \mu \in \mathbf{X}_{\bar{z}}$ if $\mu \in \mathbf{X}_{\bar{z}}$ since mass is conserved and the second moment grows linearly in time along the heat flow. Hence, by part (a), $\mathcal{H}(\mu_i) - \mathcal{H}(\mathbf{G}_\delta * \mu_i)$ is δ -bounded which yields the claim.
(B) $\mathcal{H}(\mu_i) - \mathcal{H}(\mathbf{G}_\delta * \mu_i) \leq \mathcal{H}(\mu_i) < \infty$ by nonnegativity and part (a).
(c) This is obvious thanks to smoothness and concavity of the \mathbf{m}_j ; take $L = \max_j \mathbf{m}'_j(S_j^\ell)$.
(d) This follows by means of assumption (F) on f and Taylor's theorem, $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and the fact that $\mathcal{M}(\mathbb{R}; S) \subset L^\infty(\mathbb{R}; \mathbb{R}^n)$. Note that in both cases, $\mathbf{X}_{\bar{z}} \subset L^2(\mathbb{R}; \mathbb{R}^n)$ holds.
(e) Thanks to convexity and nonnegativity of f , this is clear.

□

6.2. Time-discrete solution. We construct a time-discrete solution by means of the minimizing movement scheme (cf. Section 2) and introduce the *Yosida penalized* entropy \mathcal{E} , i.e.

$$\mathcal{E}_\tau : \mathcal{M}(\mathbb{R}; S) \times \mathcal{M}(\mathbb{R}; S) \rightarrow \mathbb{R} \cup \{\infty\}, \quad \mathcal{E}_\tau(\mu | \tilde{\mu}) := \frac{1}{2\tau} \mathbf{W}_{\mathbf{M}}(\mu, \tilde{\mu})^2 + \mathcal{E}(\mu),$$

where $\tau \in (0, \bar{\tau}]$ is a given step size; and $\bar{\tau} > 0$.

Proposition 6.4 (Minimizing movement (A)+(B)). *Let $\tau > 0$ and $\tilde{\mu} \in \mathbf{X}_{\bar{z}}$. Then, there exists a minimizer $\mu^* \in \mathbf{X}_{\bar{z}}$ of the functional $\mathcal{E}_\tau(\cdot | \tilde{\mu})$ on $\mathcal{M}(\mathbb{R}; S)$. Moreover, one has*

$$\tau \|\partial_x \mu^*\|_{L^2}^2 \leq \frac{2}{C_f} [\mathcal{H}(\tilde{\mu}) - \mathcal{H}(\mu^*)] + C\tau, \quad (49)$$

where $C = C(f, \eta) > 0$. In particular, $\mu^* - \bar{z} \in H^1(\mathbb{R}; \mathbb{R}^n)$.

Proof. By Proposition 6.3(d), \mathcal{E} is bounded from below. Hence, $\mathcal{E}_\tau(\cdot | \tilde{\mu})$ is proper and bounded from below. An infimizing sequence $(\mu_k)_{k \in \mathbb{N}}$ in $\mathbf{X}_{\bar{z}}$,

$$\lim_{k \rightarrow \infty} \mathcal{E}_\tau(\mu_k | \tilde{\mu}) = \inf \mathcal{E}_\tau(\cdot | \tilde{\mu}),$$

thus satisfies $\|\mu_k - \bar{z}\|_{L^2} \leq C$ (thanks to (F) in case (B); for case (A), this is trivial because of the uniform L^1 and L^∞ bound on $\mu_k - \bar{z}$) and $\mathbf{W}_{\mathbf{M}}(\mu_k, \tilde{\mu}) \leq C$ for some constant $C > 0$. Using Proposition 4.12(b) and Alaoglu's theorem yields the existence of a (non-relabelled) subsequence and a limit $\mu^* \in \mathbf{X}_{\bar{z}}$ such that

$\mu_k - \bar{z} \rightharpoonup \mu^* - \bar{z}$ weakly in $L^2(\mathbb{R}; \mathbb{R}^n)$ and $\mu_k \xrightarrow{*} \mu^*$ weakly* in $\mathcal{M}(\mathbb{R}; S)$, as $k \rightarrow \infty$. Note that in case (A), finiteness of $\ell_2(\mu^* - \bar{z})$ is a consequence of the uniform bound

$$\ell_2(\mu_k - \bar{z}) \leq e^L(\ell_2(\tilde{\mu} - \bar{z}) + C^2) < \infty,$$

using Proposition 4.15. The lower semicontinuity properties from the Propositions 4.12(a) and 6.3(e) show that μ^* is indeed a minimizer of $\mathcal{E}_\tau(\cdot | \tilde{\mu})$.

In order to obtain (49), recall that the heat entropy \mathcal{H} is 0-geodesically convex w.r.t. \mathbf{W}_M , thanks to Proposition 5.2. Application of the flow interchange lemma 2.3 yields

$$\tau D^{\mathcal{H}} \mathcal{E}(\mu^*) \leq \mathcal{H}(\tilde{\mu}) - \mathcal{H}(\mu^*). \quad (50)$$

For the dissipation, we obtain (write $\mu_s := S_s^{\mathcal{H}}(\mu^*)$ for brevity) for small $s > 0$:

$$\begin{aligned} -\frac{d}{ds} \mathcal{E}(\mu_s) &= -\int_{\mathbb{R}} (\nabla_z f(\mu_s) - \nabla_z f(\bar{z}) + \eta)^T \partial_{xx} \mu_s \, dx = \int_{\mathbb{R}} (\partial_x \mu_s^T \nabla_z^2 f(\mu_s) \partial_x \mu_s + \partial_x \eta^T \partial_x \mu_s) \, dx \\ &\geq \int_{\mathbb{R}} \left[C_f |\partial_x \mu_s|^2 - \frac{1}{2C_f} |\partial_x \eta|^2 - \frac{C_f}{2} |\partial_x \mu_s|^2 \right] \, dx = \frac{C_f}{2} \|\partial_x \mu_s\|_{L^2}^2 - \tilde{C}, \end{aligned}$$

where we used (F), the Cauchy-Schwarz and the Young inequality. Note that since $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$, $\tilde{C} = \tilde{C}(f, \eta)$ is finite. Passing to $s \searrow 0$ yields thanks to lower semicontinuity of the right-hand side

$$D^{\mathcal{H}} \mathcal{E}(\mu^*) \geq \frac{C_f}{2} \|\partial_x \mu^*\|_{L^2}^2 - \tilde{C},$$

from which (49) follows by insertion into (50). \square

The scheme (10) is well-posed and produces a sequence $(\mu_\tau^k)_{k \in \mathbb{N}}$ for each initial datum $\mu_\tau^0 = \mu^0 \in \mathbf{X}_{\bar{z}}$. We define the *time-discrete solution* $\mu_\tau : [0, \infty) \rightarrow \mathbf{X}_{\bar{z}}$ by piecewise constant interpolation as in (11).

The following statements are an immediate consequence of the minimizing movement:

Proposition 6.5 (Classical estimates (A)+(B)). *The following statements hold:*

- (a) For all $k \in \mathbb{N}$, one has $\mathcal{E}(\mu_\tau^k) \leq \mathcal{E}(\mu^0) < \infty$.
- (b) $\sum_{k=1}^{\infty} \mathbf{W}_M^2(\mu_\tau^k, \mu_\tau^{k-1}) \leq 2\tau(\mathcal{E}(\mu^0) - \inf \mathcal{E})$.
- (c) For all $T > 0$ and all $s, t \in [0, T]$, one has

$$\mathbf{W}_M(\mu_\tau(s), \mu_\tau(t)) \leq [2(\mathcal{E}(\mu^0) - \inf \mathcal{E}) \max(\tau, |t - s|)]^{1/2}.$$

Proof. This is classical, see for instance [3, Ch. 3]. \square

For clarity, we introduce the following notation for a given function $\varphi : [0, \infty) \rightarrow \mathbb{R}$: For each $\tau > 0$ and $s \geq 0$, let

$$\varphi_\tau(s) := \varphi\left(\left\lfloor \frac{s}{\tau} \right\rfloor \tau\right).$$

Lemma 6.6 (Discrete weak formulation (A)+(B)). *Let $\alpha > 0$, $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$, $\psi \in C_c^\infty((0, \infty)) \cap C^0([0, \infty))$ and set $\lambda = \lambda(\alpha) = -C(\frac{1}{\alpha} + 1)$ with C from Proposition 5.8(b). Then, the discrete solution μ_τ obtained from the scheme (10) satisfies the following discrete weak formulation:*

$$\begin{aligned} &\left| \int_0^\infty \int_{\mathbb{R}} \left[\rho^T \mu_\tau \frac{\psi_\tau(t) - \psi_\tau(t + \tau)}{\tau} + \psi_\tau(t) [\partial_x \rho^T \mathbf{M}(\mu_\tau) \nabla_z^2 f(\mu_\tau) \partial_x \mu_\tau + \partial_x \rho^T \mathbf{M}(\mu_\tau) \partial_x \eta] \right] \, dx \, dt \right| \\ &\leq \left| \alpha \int_0^\infty \int_{\mathbb{R}} \left[h_{\bar{z}}(\mu_\tau) \frac{|\psi|_\tau(t) - |\psi|_\tau(t + \tau)}{\tau} + |\psi|_\tau(t) [\partial_x \mu_\tau^T \nabla_z^2 f(\mu_\tau) \partial_x \mu_\tau + \partial_x \eta^T \partial_x \mu_\tau] \right] \, dx \, dt \right| \quad (51) \\ &\quad + 2\lambda\tau \|\psi\|_{C^0} [\mathcal{E}(\mu^0) - \inf \mathcal{E}]. \end{aligned}$$

Proof. Recall that for this choice of λ , the regularized potential energy \mathcal{V} defined in (45) is λ -geodesically convex w.r.t. \mathbf{W}_M (cf. Proposition 5.8). Hence, we are in position to apply the flow interchange lemma 2.3 to obtain for all $k \in \mathbb{N}$:

$$\mathcal{V}(\mu_\tau^k) + \tau D^\mathcal{V} \mathcal{E}(\mu_\tau^k) + \frac{\lambda}{2} \mathbf{W}_M^2(\mu_\tau^k, \mu_\tau^{k-1}) \leq \mathcal{V}(\mu_\tau^{k-1}). \quad (52)$$

For the dissipation, one has (write $\mu_s := S_s^\mathcal{V}(\mu_\tau^k)$ for brevity) for small $s > 0$

$$\begin{aligned} -\frac{d}{ds} \mathcal{E}(\mu_s) &= -\int_{\mathbb{R}} [\nabla_z f(\mu_s) - \nabla_z f(\bar{z}) + \eta]^T [\alpha \partial_{xx} \mu_s + \partial_x (\mathbf{M}(\mu_s) \partial_x \rho)] dx \\ &= \alpha \int_{\mathbb{R}} [\partial_x \mu_s^T \nabla_z^2 f(\mu_s) \partial_x \mu_s + \partial_x \eta^T \partial_x \mu_s] dx + \int_{\mathbb{R}} [\partial_x \rho^T \mathbf{M}(\mu_s) \nabla_z^2 f(\mu_s) \partial_x \mu_s + \partial_x \rho^T \mathbf{M}(\mu_s) \partial_x \eta] dx, \end{aligned}$$

and consequently, passing to $s \searrow 0$:

$$D^\mathcal{V} \mathcal{E}(\mu_\tau^k) \geq \alpha \int_{\mathbb{R}} [\partial_x \mu_\tau^k{}^T \nabla_z^2 f(\mu_\tau^k) \partial_x \mu_\tau^k + \partial_x \eta^T \partial_x \mu_\tau^k] dx + \int_{\mathbb{R}} [\partial_x \rho^T \mathbf{M}(\mu_\tau^k) \nabla_z^2 f(\mu_\tau^k) \partial_x \mu_\tau^k + \partial_x \rho^T \mathbf{M}(\mu_\tau^k) \partial_x \eta] dx.$$

Inserting this into (52) and repeating this calculation with $-\rho$ in place of ρ yields

$$\begin{aligned} &\left| \int_{\mathbb{R}} [\rho^T [\mu_\tau^k - \mu_\tau^{k-1}] + \tau \partial_x \rho^T \mathbf{M}(\mu_\tau^k) \nabla_z^2 f(\mu_\tau^k) \partial_x \mu_\tau^k + \tau \partial_x \rho^T \mathbf{M}(\mu_\tau^k) \partial_x \eta] dx \right| \\ &\leq \left| \alpha \int_{\mathbb{R}} [h_{\bar{z}}(\mu_\tau^k) - h_{\bar{z}}(\mu_\tau^{k-1}) + \tau \partial_x \mu_\tau^k{}^T \nabla_z^2 f(\mu_\tau^k) \partial_x \mu_\tau^k + \tau \partial_x \eta^T \partial_x \mu_\tau^k] dx + \frac{\lambda}{2} \mathbf{W}_M^2(\mu_\tau^k, \mu_\tau^{k-1}) \right|. \end{aligned} \quad (53)$$

Let $\psi \in C^0([0, \infty))$ be *nonnegative* and have compact support in $(0, \infty)$. We multiply the chain of inequalities (53) with $\psi((k-1)\tau)$ and take the sum over all $k \in \mathbb{N}$, recalling Proposition 6.5(b) and observing

$$\sum_{k \in \mathbb{N}} \psi((k-1)\tau) [g(\mu_\tau^k) - g(\mu_\tau^{k-1})] = \sum_{k \in \mathbb{N}} g(\mu_\tau^k) [\psi((k-1)\tau) - \psi(k\tau)],$$

for an arbitrary map $g : \mathbb{R}^n \rightarrow \mathbb{R}^d$, $d \in \mathbb{N}$. The resulting chain of inequalities can be expressed with the discrete solution μ_τ as follows:

$$\begin{aligned} &\left| \int_0^\infty \int_{\mathbb{R}} \left[\rho^T \mu_\tau \frac{\psi_\tau(t) - \psi_\tau(t+\tau)}{\tau} + \psi_\tau(t) [\partial_x \rho^T \mathbf{M}(\mu_\tau) \nabla_z^2 f(\mu_\tau) \partial_x \mu_\tau + \partial_x \rho^T \mathbf{M}(\mu_\tau) \partial_x \eta] \right] dx dt \right| \\ &\leq \left| \alpha \int_0^\infty \int_{\mathbb{R}} \left[h_{\bar{z}}(\mu_\tau) \frac{\psi_\tau(t) - \psi_\tau(t+\tau)}{\tau} + \psi_\tau(t) [\partial_x \mu_\tau^T \nabla_z^2 f(\mu_\tau) \partial_x \mu_\tau + \partial_x \eta^T \partial_x \mu_\tau] \right] dx dt \right. \\ &\quad \left. + \lambda \tau \|\psi\|_{C^0} [\mathcal{E}(\mu^0) - \inf \mathcal{E}] \right|. \end{aligned} \quad (54)$$

For general $\psi \in C_c^\infty((0, \infty)) \cap C^0([0, \infty))$, decompose ψ into its positive and negative part and subtract the respective inequalities (54) to obtain (51). \square

6.3. Passage to continuous time.

Proposition 6.7 (*A priori estimates (A)*). *For given $T > 0$, there exist constants $C_i = C_i(T) > 0$ such that for all $\tau \in (0, \bar{\tau}]$, the following holds:*

- (a) $\mathbf{W}_M(\mu_\tau(t), \mu^0) \leq C_1$ for all $t \in [0, T]$.
- (b) $\|\mu_\tau - \bar{z}\|_{L^\infty([0, T]; L^2)} \leq C_2$.
- (c) $\ell_2(\mu_\tau(t) - \bar{z}) \leq C_3$ for all $t \in [0, T]$.
- (d) $\|\mu_\tau - \bar{z}\|_{L^2([0, T]; H^1)} \leq C_4$.

Proof. (a) Using Proposition 6.5(c) yields

$$\mathbf{W}_M(\mu_\tau(t), \mu^0) \leq [2(\mathcal{E}(\mu^0) - \inf \mathcal{E}) \max(\tau, t)]^{1/2} \leq C_1$$

for $0 \leq t \leq T$ and $0 < \tau \leq \bar{\tau}$.

(b) This is obvious thanks to the uniform bounds on $\mu_\tau - \bar{z}$ in $L^1(\mathbb{R}; \mathbb{R}^n)$ and $L^\infty(\mathbb{R}; \mathbb{R}^n)$, respectively.

(c) By part (a) and Proposition 4.15, one has

$$\ell_2(\mu_\tau(t) - \bar{z}) \leq e^L(\ell_2(\mu^0 - \bar{z}) + C_1^2) \quad \text{for all } t \in [0, T].$$

(d) In view of (b), it remains to prove that $\|\partial_x \mu_\tau\|_{L^2([0, T]; L^2)}$ is τ -uniformly bounded. Set $K := \lfloor \frac{T}{\tau} \rfloor + 1 \leq \frac{T + \bar{\tau}}{\tau}$ to obtain

$$\int_0^T \|\partial_x \mu_\tau(t)\|_{L^2}^2 dt \leq \sum_{k=1}^K \tau \|\partial_x \mu_\tau^k\|_{L^2}^2 \leq \sum_{k=1}^K \left[\frac{2}{C_f} (\mathcal{H}(\mu_\tau^{k-1}) - \mathcal{H}(\mu_\tau^k)) + C\tau \right], \quad (55)$$

where we used (49) in the last step. In the proof of Proposition 6.3(a), we have seen that there exist constants $\tilde{C}_0, \tilde{C}_1 > 0$ such that for all $\mu \in \mathbf{X}_{\bar{z}}$

$$|\mathcal{H}(\mu)| \leq \tilde{C}_0 + \tilde{C}_1(m + \ell_2(\mu - \bar{z})).$$

Using (c) with $T + \bar{\tau}$ in place of T , we eventually end up with

$$\int_0^T \|\partial_x \mu_\tau(t)\|_{L^2}^2 dt \leq C(T + \bar{\tau}) + \frac{2}{C_f} (\mathcal{H}(\mu^0) + \tilde{C}_0 + \tilde{C}_1(m + C_3)). \quad \square$$

Proposition 6.8 (*A priori estimates (B)*). *For given $T > 0$, there exist constants $C_i = C_i(T) > 0$ such that for all $\tau \in (0, \bar{\tau}]$, the following holds:*

- (a) $\mathbf{W}_M(\mu_\tau(t), \mu^0) \leq C_1$ for all $t \in [0, T]$.
- (b) $\|\mu_\tau - \bar{z}\|_{L^\infty([0, T]; L^2)} \leq C_2$.
- (d) $\|\mu_\tau - \bar{z}\|_{L^2([0, T]; H^1)} \leq C_4$.

Proof. Part (a) is the same as for Proposition 6.7. For part (b), thanks to Proposition 6.5(a), for all $t > 0$, one has $\mathcal{E}(\mu_\tau(t)) = \mathcal{E}(\mu_\tau^k) \leq \mathcal{E}(\mu^0)$, with $k = \lfloor \frac{t}{\tau} \rfloor + 1$. Using Proposition 6.3(d) yields

$$\|\mu_\tau(t) - \bar{z}\|_{L^2} \leq C_2 \quad \text{for all } t > 0.$$

For (d), we again proceed as before to arrive at (55). From there, the claim obviously follows by nonnegativity of \mathcal{H} . \square

We now are in position to pass to the limit $\tau \searrow 0$.

Proposition 6.9 (Continuous-time limit (A)+(B)). *Let $T > 0$ be given, $(\tau_k)_{k \in \mathbb{N}}$ be a vanishing sequence of step sizes, i.e. $\tau_k \searrow 0$ as $k \rightarrow \infty$, and $(\mu_{\tau_k})_{k \in \mathbb{N}}$ be the corresponding sequence of discrete solutions obtained by the minimizing movement scheme. Then, there exists a (non-relabelled) subsequence and a limit curve $\mu : [0, T] \rightarrow \mathbf{X}_{\bar{z}}$ such that as $k \rightarrow \infty$:*

- (a) For fixed $t \in [0, T]$, $\mu_{\tau_k}(t) \xrightarrow{*} \mu(t)$ weakly* in $\mathcal{M}(\mathbb{R}; S)$,
- (b) $\mu_{\tau_k} - \bar{z} \rightharpoonup \mu - \bar{z}$ weakly in $L^2([0, T]; H^1(\mathbb{R}; \mathbb{R}^n))$,
- (c) $\mu_{\tau_k} - \bar{z} \rightarrow \mu - \bar{z}$ strongly in $L^2([0, T]; L_{\text{loc}}^2(\mathbb{R}; \mathbb{R}^n))$,

with the properties

$$\mu \in C^{1/2}([0, T]; (\mathcal{M}(\mathbb{R}; S), \mathbf{W}_M)), \quad (56)$$

$$\mu - \bar{z} \in L^\infty([0, T]; L^2(\mathbb{R}; \mathbb{R}^n)) \cap L^2([0, T]; H^1(\mathbb{R}; \mathbb{R}^n)). \quad (57)$$

Moreover, the limit μ is a weak solution to (1) in the following sense: For all $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and all $\psi \in C_c^\infty((0, \infty)) \cap C^0([0, \infty))$, one has

$$\int_0^\infty \int_{\mathbb{R}} [-\partial_t \psi \rho^T \mu + \psi [\partial_x \rho^T \mathbf{M}(\mu) \nabla_z^2 f(\mu) \partial_x \mu + \partial_x \rho^T \mathbf{M}(\mu) \partial_x \eta]] dx dt = 0. \quad (58)$$

Proof. We divide the proof into several steps.

Step 1: Weak convergence and limit properties.

Using the *a priori* estimates in Proposition 6.7/6.8(a)&(b) together with Proposition 4.12 and Alaoglu's theorem, we deduce the weak convergences (a)&(b) and also the properties of the limit. Note that in case (A), finiteness of $\ell_2(\mu(t) - \bar{z})$ is a consequence of the uniform estimate from Proposition 6.7(c). In both cases, 1/2-Hölder continuity w.r.t. \mathbf{W}_M can be obtained thanks to Proposition 6.5(c) via a refined version of the Arzelà-Ascoli theorem (cf. [3, Thm. 3.3.1]).

Step 2: Strong convergence.

In order to prove the strong convergence (c), we fix a bounded interval $I \subset \mathbb{R}$ and apply Theorem 2.4 with the admissible choices

$$\mathbf{Y} := \{u \in \mathcal{M}(I; S) : u - \bar{z} \in L^2(I; \mathbb{R}^n)\}, \text{ endowed with } \|u\|_{\mathbf{Y}} := \|u - \bar{z}\|_{L^2(I)},$$

which is isometric to a closed subset of $L^2(I; \mathbb{R}^n)$,

$$\mathcal{A}(u) := \begin{cases} \|u - \bar{z}\|_{H^1(I)}^2, & \text{if } u - \bar{z} \in H^1(I; \mathbb{R}^n) \\ +\infty, & \text{otherwise,} \end{cases}$$

which has relatively compact sublevels in \mathbf{Y} due to the Rellich-Kondrachov compactness theorem, and

$$\mathbf{W}(u, \tilde{u}) := \begin{cases} \mathbf{W}_{\mathbf{M}}(u_{\bar{z}}, (\tilde{u})_{\bar{z}}), & \text{if } u, \tilde{u} \in \text{Dom}(\mathcal{A}), \\ +\infty, & \text{otherwise,} \end{cases}$$

where $u_{\bar{z}}$ is to be understood as the extension of $u \in \mathcal{M}(I; S)$ to a function in $\mathcal{M}(\mathbb{R}; S)$ by setting $u_{\bar{z}} \equiv \bar{z}$ outside I . We verify the hypotheses (12)&(13) for the sequence $U := (\mu_{\tau_k})_{k \in \mathbb{N}}$, where – for the sake of clarity – we identify μ_{τ_k} with its spatial restriction $\mu_{\tau_k}|_{[0, \infty) \times I}$:

(12) is immediate because of the *a priori* estimate from Proposition 6.7/6.8(d). For (13), we claim

$$\sup_{k \in \mathbb{N}} \int_0^{T-h} \mathbf{W}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t)) dt \leq \max(1, \sqrt{T+\bar{\tau}}) \sqrt{2(\mathcal{E}(\mu^0) - \inf \mathcal{E})(T+\bar{\tau})h} \quad \text{for all } h \in (0, \bar{\tau}), \quad (59)$$

from which (13) follows. Indeed, for fixed $k \in \mathbb{N}$ and $h \in (0, \tau_k]$, one has

$$\begin{aligned} \int_0^{T-h} \mathbf{W}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t)) dt &= \sum_{i=1}^{\lfloor \frac{T}{\tau_k} \rfloor} h \mathbf{W}_{\mathbf{M}}(\mu_{\tau_k}^i, \mu_{\tau_k}^{i+1}) \leq \sqrt{2(\mathcal{E}(\mu^0) - \inf \mathcal{E})} \sqrt{h^2 \left\lfloor \frac{T}{\tau_k} \right\rfloor} \\ &\leq \sqrt{2(\mathcal{E}(\mu^0) - \inf \mathcal{E})(T+\bar{\tau})h}, \end{aligned}$$

thanks to Hölder's inequality and Proposition 6.5(b). On the other hand, for $h \in (\tau_k, \bar{\tau}]$, we directly get from Proposition 6.5(c):

$$\int_0^{T-h} \mathbf{W}(\mu_{\tau_k}(t+h), \mu_{\tau_k}(t)) dt \leq (T-h) \sqrt{2(\mathcal{E}(\mu^0) - \inf \mathcal{E})h} \leq (T+\bar{\tau}) \sqrt{2(\mathcal{E}(\mu^0) - \inf \mathcal{E})h}.$$

Hence, (59) holds and the application of Theorem 2.4 yields the existence of a (non-relabelled) subsequence which converges to (the spatial restriction to I of) μ in measure w.r.t. $t \in (0, T)$. By the uniform estimate in Proposition 6.7/6.8(b) and the dominated convergence theorem, we conclude that

$$\mu_{\tau_k} - \bar{z} \rightarrow \mu - \bar{z} \text{ strongly in } L^2([0, T] \times I; \mathbb{R}^n),$$

proving claim (c) for a prescribed interval I . By a diagonal argument, setting $I_R := [-R, R]$ and letting $R \nearrow \infty$, we deduce that (c) is true simultaneously for every bounded interval I , extracting a further subsequence. Moreover, we may assume that μ_{τ_k} converges to μ pointwise almost everywhere in $[0, T] \times \mathbb{R}$.

Step 3: Weak formulation.

Let $\rho \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and $\psi \in C_c^\infty((0, \infty)) \cap C^0([0, \infty))$ be given and set $\alpha_k := \sqrt{\tau_k}$ for $k \in \mathbb{N}$. By Lemma 6.6, μ_{τ_k} satisfies the discrete weak formulation (51) for each k , putting $\lambda_k = \lambda(\alpha_k)$ according to Lemma 6.6. Note that with this choice of α_k , one has $\lim_{k \rightarrow \infty} \lambda_k \tau_k = 0$.

We first prove that

$$\int_0^\infty \int_{\mathbb{R}} \left[h_{\bar{z}}(\mu_{\tau_k}) \frac{|\psi|_{\tau_k}(t) - |\psi|_{\tau_k}(t+\tau_k)}{\tau_k} + |\psi|_{\tau_k}(t) [\partial_x \mu_{\tau_k}^T \nabla_z^2 f(\mu_{\tau_k}) \partial_x \mu_{\tau_k} + \partial_x \eta^T \partial_x \mu_{\tau_k}] \right] dx dt \quad (60)$$

is bounded w.r.t. $k \in \mathbb{N}$. For the first part, since $\psi \in C_c^\infty((0, \infty))$, there exists $T' > 0$ such that

$$\left| \int_0^\infty \int_{\mathbb{R}} h_{\bar{z}}(\mu_{\tau_k}) \frac{|\psi|_{\tau_k}(t) - |\psi|_{\tau_k}(t+\tau_k)}{\tau_k} dx dt \right| \leq C \int_0^{T'} \int_{\mathbb{R}} |h_{\bar{z}}(\mu_{\tau_k})| dx dt.$$

In case (A), we obtain

$$\int_0^{T'} \int_{\mathbb{R}} |h_{\bar{z}}(\mu_{\tau_k})| dx dt \leq \int_0^{T'} [\tilde{C}_0 + \tilde{C}_1(m + \ell_2(\mu_{\tau_k}(t) - \bar{z}))] dt,$$

which is bounded thanks to Proposition 6.7(c). In case (B), we have

$$\int_0^{T'} \int_{\mathbb{R}} |h_{\bar{z}}(\mu_{\tau_k})| dx dt \leq \tilde{C} \int_0^{T'} \|\mu_{\tau_k}(t) - \bar{z}\|_{L^2}^2 dt,$$

so Proposition 6.8(b) yields boundedness. For the second part in (60), we use $\mathcal{M}(\mathbb{R}; S) \subset L^\infty(\mathbb{R}; \mathbb{R}^n)$ and the Cauchy-Schwarz inequality to obtain

$$\left| \int_0^\infty \int_{\mathbb{R}} |\psi|_{\tau_k}(t) [\partial_x \mu_{\tau_k}^\top \nabla_z^2 f(\mu_{\tau_k}) \partial_x \mu_{\tau_k} + \partial_x \eta^\top \partial_x \mu_{\tau_k}] dx dt \right| \leq \overline{C} \int_0^{T'} \int_{\mathbb{R}} \left[\overline{C}' |\partial_x \mu_{\tau_k}|^2 + \frac{1}{2} |\partial_x \eta|^2 \right] dx dt.$$

Thanks to $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$ and Proposition 6.7/6.8(d), this is bounded.

From the dominated convergence theorem, since μ_{τ_k} converges to μ pointwise a.e., it follows – using $\mathcal{M}(\mathbb{R}; S) \subset L^\infty(\mathbb{R}; \mathbb{R}^n)$ again – that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \left(\int_0^\infty \int_{\mathbb{R}} \left[\rho^\top \mu_{\tau_k} \frac{\psi_{\tau_k}(t) - \psi_{\tau_k}(t + \tau_k)}{\tau_k} + \psi_{\tau_k}(t) \partial_x \rho^\top \mathbf{M}(\mu_{\tau_k}) \partial_x \eta \right] dx dt \right) \\ &= \int_0^\infty \int_{\mathbb{R}} [-\partial_t \psi \rho^\top \mu + \psi \partial_x \rho^\top \mathbf{M}(\mu) \partial_x \eta] dx dt. \end{aligned} \quad (61)$$

We now prove

$$\lim_{k \rightarrow \infty} \left(\int_0^\infty \int_{\mathbb{R}} \psi_{\tau_k}(t) \partial_x \rho^\top \mathbf{M}(\mu_{\tau_k}) \nabla_z^2 f(\mu_{\tau_k}) \partial_x \mu_{\tau_k} dx dt \right) = \int_0^\infty \int_{\mathbb{R}} \psi \partial_x \rho^\top \mathbf{M}(\mu) \nabla_z^2 f(\mu) \partial_x \mu dx dt. \quad (62)$$

First, we show that

$$\lim_{k \rightarrow \infty} \int_0^\infty \int_{\mathbb{R}} (\psi_{\tau_k}(t) \partial_x \rho^\top \mathbf{M}(\mu_{\tau_k}) \nabla_z^2 f(\mu_{\tau_k}) - \psi \partial_x \rho^\top \mathbf{M}(\mu) \nabla_z^2 f(\mu)) \partial_x \mu_{\tau_k} dx dt = 0. \quad (63)$$

Using Hölder's inequality, the fact that ψ and ρ have compact support and Proposition 6.7/6.8(d) reduces the problem to verifying

$$\lim_{k \rightarrow \infty} \int_0^\infty \int_{\mathbb{R}} |\psi_{\tau_k}(t) \partial_x \rho^\top \mathbf{M}(\mu_{\tau_k}) \nabla_z^2 f(\mu_{\tau_k}) - \psi \partial_x \rho^\top \mathbf{M}(\mu) \nabla_z^2 f(\mu)|^2 dx dt = 0.$$

We proceed using dominated convergence since the integrand converges pointwise a.e. to zero and the following pointwise estimate holds:

$$|\psi_{\tau_k}(t) \partial_x \rho^\top \mathbf{M}(\mu_{\tau_k}) \nabla_z^2 f(\mu_{\tau_k}) - \psi \partial_x \rho^\top \mathbf{M}(\mu) \nabla_z^2 f(\mu)|^2 \leq C \mathbf{1}_{\text{supp } \psi} \mathbf{1}_{\text{supp } \rho}.$$

The r.h.s. obviously is integrable on $(0, \infty) \times \mathbb{R}$.

Second,

$$\lim_{k \rightarrow \infty} \int_0^\infty \int_{\mathbb{R}} \psi \partial_x \rho^\top \mathbf{M}(\mu) \nabla_z^2 f(\mu) (\partial_x \mu_{\tau_k} - \partial_x \mu) dx dt = 0, \quad (64)$$

since $\psi \partial_x \rho^\top \mathbf{M}(\mu) \nabla_z^2 f(\mu)$ is bounded and has compact support in $[0, T'] \times \mathbb{R}$ for some $T' > 0$ and hence is an element of $L^2([0, T']; L^2(\mathbb{R}; \mathbb{R}^n))$, yielding the claim together with $\partial_x \mu_{\tau_k} \rightharpoonup \partial_x \mu$ weakly in $L^2([0, T']; L^2(\mathbb{R}; \mathbb{R}^n))$ by part (b) of this proposition. We have thus proved (62). Putting (60)–(62) together yields (58). \square

We summarize the results of this section in the following

Theorem 6.10 (Existence of weak solutions). *Consider the initial-value problem for the system of degenerate diffusion equations with drift*

$$\partial_t \mu = \partial_x (\mathbf{M}(\mu) \nabla_z^2 f(\mu) \partial_x \mu + \mathbf{M}(\mu) \partial_x \eta), \quad t > 0, x \in \mathbb{R}, \quad (65)$$

$$\mu(0, x) = \mu^0(x), \quad x \in \mathbb{R}, \quad (66)$$

where the mobility \mathbf{M} is fully decoupled on the state space $S = [S^\ell, S^r] \subset \mathbb{R}^n$ and of the form $\mathbf{M}(z) = (\nabla_z^2 h(z))^{-1} \in \mathbb{R}^{n \times n}$ with $h : S \rightarrow \mathbb{R}$ satisfying (H0)–(H3). Assume that $f : S \rightarrow \mathbb{R}$ satisfies (F) and $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}^n)$.

Suppose that $\mu^0 \in \mathcal{M}(\mathbb{R}; S)$ and either

(A) $\mu^0 - \bar{z} \in L^1(\mathbb{R}; \mathbb{R}^n)$ and $\ell_2(\mu^0 - \bar{z}) < \infty$ for $\bar{z} := S^\ell$

or

(B) $\mu^0 - \bar{z} \in L^2(\mathbb{R}; \mathbb{R}^n)$ for some $\bar{z} \in \text{int}(S)$.

Then, there exists a function $\mu : [0, \infty) \times \mathbb{R} \rightarrow S$ with

$$\begin{aligned} \mu &\in C^{1/2}([0, T]; (\mathcal{M}(\mathbb{R}; S), \mathbf{W}_M)), \\ \mu - \bar{z} &\in L^\infty([0, T]; L^2(\mathbb{R}; \mathbb{R}^n)) \cap L^2([0, T]; H^1(\mathbb{R}; \mathbb{R}^n)) \end{aligned}$$

for all $T > 0$ satisfying (65) in the sense of distributions and attaining the initial condition (66). Additionally, in case (A), the following holds for all $t \in [0, T]$:

$$\|\mu(t) - \bar{z}\|_{L^1} = \|\mu^0 - \bar{z}\|_{L^1}, \quad \text{and} \quad \ell_2(\mu(t) - \bar{z}) < \infty.$$

APPENDIX A. PROOF OF PROPOSITION 5.8(a)

Since in the case at hand, the system (46) is decoupled, it suffices to prove the assertion in the scalar case $n = 1$, where the mobility \mathbf{m} is a scalar function satisfying the properties of Section 5.4.2. Suppose that $\mu^0 \in \mathcal{M}(\mathbb{R}; S)$ attains values in $\text{int}(S)$ only, with $S = [S^\ell, S^r] \subset \mathbb{R}$ being an interval, and $\mu^0 - \bar{z} \in H^1(\mathbb{R})$ for some $\bar{z} \in S$. Using the transformation $u := \mu - \bar{z}$ and writing $\theta := \rho_x$, we may instead consider the equation

$$\partial_t u = \partial_{xx} u + \partial_x(\mathbf{m}(u + \bar{z})\theta), \quad (67)$$

together with the initial condition $u^0 := \mu^0 - \bar{z} \in H^1(\mathbb{R})$ with values in $(S^\ell - \bar{z}, S^r - \bar{z}) \ni 0$.

Inspired from [16, Ch. 3], we write (67) as an abstract semilinear evolution equation on $H^1(\mathbb{R})$:

$$\dot{u}(t) = -Au(t) + F(u(t)), \quad (68)$$

with $A := -\frac{d^2}{dx^2}$, and

$$F(u) := \mathbf{m}'(u + \bar{z})u_x\theta + \mathbf{m}(u + \bar{z})\theta_x.$$

We first prove some properties of the nonlinearity F .

Lemma A.1 (Properties of F). (a) F maps bounded subsets of $H^1(\mathbb{R})$ onto bounded subsets of $L^2(\mathbb{R})$, because for all $u \in H^1(\mathbb{R})$, one has

$$\|F(u)\|_{L^2} \leq C_0\|u\|_{H^1} + C_1, \quad (69)$$

for some $C_0, C_1 > 0$.

(b) F is locally Lipschitz continuous in the following sense: If $u, v \in H^1(\mathbb{R})$ with $\|u - u^0\|_{H^1} < \delta$ and $\|v - u^0\|_{H^1} < \delta$ for some $\delta > 0$, then

$$\|F(u) - F(v)\|_{L^2} \leq C_2\|u - v\|_{H^1}, \quad (70)$$

for some $C_2 = C_2(\delta, u^0) > 0$.

Proof. (a) By the triangle inequality, we have

$$\begin{aligned} \|F(u)\|_{L^2} &\leq \|\mathbf{m}\|_{C^1}\|\theta\|_{C^0}\|u_x\|_{L^2} + \|[\mathbf{m}(u + \bar{z}) - \mathbf{m}(\bar{z})]\theta_x\|_{L^2} + \|\mathbf{m}(\bar{z})\theta_x\|_{L^2} \\ &\leq 2\|\mathbf{m}\|_{C^1}\|\theta\|_{C^1}\|u\|_{H^1} + \|\mathbf{m}(\bar{z})\theta_x\|_{L^2}, \end{aligned}$$

from which the desired estimate follows since θ has compact support.

(b) With u and v as required, one has

$$\begin{aligned} \|F(u) - F(v)\|_{L^2} &\leq \|\theta\|_{C^1}\|\mathbf{m}(u + \bar{z}) - \mathbf{m}(v + \bar{z})\|_{L^2} + \|\theta\mathbf{m}'(u + \bar{z})[u_x - v_x]\|_{L^2} \\ &\quad + \|\theta u_x^0[\mathbf{m}'(u + \bar{z}) - \mathbf{m}'(v + \bar{z})]\|_{L^2} + \|\theta(v_x - u_x^0)[\mathbf{m}'(u + \bar{z}) - \mathbf{m}'(v + \bar{z})]\|_{L^2} \\ &\leq \|\theta\|_{C^1}\|\mathbf{m}\|_{C^2}[\|u - v\|_{L^2} + \|u_x - v_x\|_{L^2}] + (\|u^0\|_{H^1} + \|v - u^0\|_{H^1})\|u - v\|_{L^\infty}. \end{aligned}$$

Since $H^1(\mathbb{R})$ is continuously embedded into $C^0(\mathbb{R})$ and $\|v - u^0\|_{H^1} < \delta$, the desired estimate follows. \square

Let now $\delta > 0$ fixed, but arbitrary and define

$$K_\delta := \{u \in C^0([0, T]; H^1(\mathbb{R})) \mid \|u(t) - u^0\|_{H^1} \leq \delta \quad \forall t \in [0, T]\},$$

where $T > 0$ is to be determined later. K_δ is a closed subset of the Banach space $C^0([0, T]; H^1(\mathbb{R}))$. Define a mapping B on K_δ by

$$B(u)(t) := e^{-At}u^0 + \int_0^t e^{-A(t-s)}F(u(s)) \, ds \quad \text{for } t \in [0, T].$$

We prove the following statement:

There exists $T = T(\delta, u^0) > 0$ sufficiently small such that B maps K_δ into itself and is a strict contraction.

We first prove that $\|B(u)(t) - u^0\|_{H^1} \leq \delta$ for all $t \in [0, T]$, where T is sufficiently small. For all $s \in (0, t)$, one has

$$e^{-A(t-s)}F(u(s)) = \mathbf{G}_{t-s} * F(u(s)),$$

where \mathbf{G} is the one-dimensional heat kernel from (28). Note that for all $\sigma > 0$, we have

$$\|\mathbf{G}_\sigma\|_{L^1} = A_0, \quad \|\partial_y \mathbf{G}_\sigma\|_{L^1} = A_1 \sigma^{-1/2}, \quad (71)$$

for some constants $A_0, A_1 > 0$. Elementary kernel estimates yield

$$\|\mathbf{G}_t * u^0 - u^0\|_{H^1} \leq \frac{\delta}{2}, \quad (72)$$

for all $t \in [0, T]$, provided that T is sufficiently small. For the other part, we use Young's inequality for convolutions to obtain

$$\left\| \int_0^t e^{-A(t-s)}F(u(s)) \, ds \right\|_{H^1} \leq \int_0^t [\|\mathbf{G}_{t-s}\|_{L^1} + \|\partial_y \mathbf{G}_{t-s}\|_{L^1}] \|F(u(s))\|_{L^2} \, ds$$

Using (71) and (69), together with the fact that $\|u(s)\|_{H^1} \leq \|u^0\|_{H^1} + \delta$, since $u \in K_\delta$ yields

$$\left\| \int_0^t e^{-A(t-s)}F(u(s)) \, ds \right\|_{H^1} \leq (tA_0 + 2\sqrt{t}A_1)(C_0\|u^0\|_{H^1} + C_0\delta + C_1) \leq \frac{\delta}{2}, \quad (73)$$

for all $t \in [0, T]$, provided that T is sufficiently small. Putting (72)&(73) together yields the claim. Along the same lines, it can be shown that $B(u) \in C^0([0, T]; H^1(\mathbb{R}))$; hence $B(u) \in K_\delta$.

For Lipschitz-continuity, we proceed exactly as before using (70) instead:

$$\begin{aligned} \|B(u)(t) - B(v)(t)\|_{H^1} &\leq C_2 \int_0^t (A_0 + (t-s)^{-1/2}A_1) \|u(t-s) - v(t-s)\|_{H^1} \, ds \\ &\leq C_2(A_0T + 2\sqrt{T}A_1) \|u - v\|_{C^0([0, T]; H^1)}, \end{aligned}$$

for all $t \in [0, T]$. Hence, if T is sufficiently small, one has

$$\|B(u) - B(v)\|_{C^0([0, T]; H^1)} \leq L \|u - v\|_{C^0([0, T]; H^1)},$$

for some $0 \leq L < 1$.

Now, by Banach's fixed point theorem, B possesses exactly one fixed point u^* in K_δ which is, by means of [16, Lemma 3.3.2], the desired unique smooth solution to (68) on $[0, T]$. It remains to prove that $u^*(t, x) \in \text{int}(S)$ for all $x \in \mathbb{R}$ and $t \in [0, T']$, for some sufficiently small $T' > 0$.

Case 1: $\bar{z} \in \text{int}(S)$. Thanks to $u^0 \in H^1(\mathbb{R})$, there exists $\delta_0 > 0$ such that

$$\text{dist}(u^0(x), \partial S) > \delta_0 \quad \forall x \in \mathbb{R}.$$

Since $u^* \in C^0([0, T]; H^1(\mathbb{R})) \subset C^0([0, T]; C^0(\mathbb{R}))$, there exists $T' \in (0, T]$ such that $\|u^*(t, \cdot) - u^0\|_{C^0} < \frac{\delta_0}{2}$ for all $t \in [0, T']$. Hence, we obtain

$$\text{dist}(u^*(t, x), \partial S) > \frac{\delta_0}{2} \quad \forall t \in [0, T'], \forall x \in \mathbb{R},$$

which proves the claim.

Case 2: $\bar{z} = S^\ell$. First, as in case 1, there exists $T'_1 \in (0, T]$ such that $u^*(t, x) < S^r - S^\ell$ for all $x \in \mathbb{R}$ and all $t \in [0, T'_1]$. It remains to prove the lower bound $u^*(t, x) > 0$. Let therefore $R > 0$ such that $\text{supp}(\theta) \subset [-R, R]$. Since u^0 is strictly positive and continuous, there exists $\delta > 0$ such that $u^0(x) > \delta$ for all $x \in [-(R+1), R+1]$. Hence, we can find $T'_2 \in (0, T'_1]$ such that $u^*(t, x) > \frac{\delta}{2}$ for all $t \in [0, T'_2]$ and all $x \in [-(R+1), R+1]$. Moreover, thanks to the smoothness of u^* , there exists $C_0 > 0$ such that $|F(u^*(s))(y)| < C_0$ for all $s \in [0, T'_2]$ and all $y \in [-R, R]$.

It remains to consider the case $|x| > R+1$, $t \in [0, T'_2]$, where we explicitly analyze u^* by means of its fixed-point property $B(u^*) = u^*$, i.e.

$$u^*(t, x) = \int_{\mathbb{R}} \mathbf{G}_t(x-y) u^0(y) dy + \int_0^t \int_{\mathbb{R}} \mathbf{G}_{t-s}(x-y) F(u^*(s))(y) dy ds. \quad (74)$$

For the second part in formula (74), we immediately obtain the estimate

$$\left| \int_0^t \int_{\mathbb{R}} \mathbf{G}_{t-s}(x-y) F(u^*(s))(y) dy ds \right| \leq C_0 \int_{-R}^R \int_0^t \mathbf{G}_s(x-y) ds dy,$$

where we recall that $|x-y| > 1$. Since for fixed $v > 0$, the map $g_v : (0, \infty) \rightarrow \mathbb{R}$, $g_v(s) := \frac{1}{\sqrt{4\pi s}} \exp\left(-\frac{v^2}{4s}\right)$ is strictly increasing for $s < \frac{v^2}{2}$, we obtain

$$\left| \int_0^t \int_{\mathbb{R}} \mathbf{G}_{t-s}(x-y) F(u^*(s))(y) dy ds \right| \leq C_0 \int_{-R}^R t \mathbf{G}_t(x-y) ds dy,$$

if $t < \frac{1}{2}$. Hence, for all $t < \min\left(T'_2, \frac{1}{2}, \frac{\delta}{2C_0}\right) =: T'$ and all $|x| > R+1$, formula (74) yields

$$u^*(t, x) > \int_{-R}^R \left(\frac{\delta}{2} - C_0 t\right) \mathbf{G}_t(x-y) dy,$$

the right-hand side being nonnegative.

Case 3: $\bar{z} = S^r$. Here, proceed in analogy to case 2. □

REFERENCES

- [1] H. Amann. Dynamic theory of quasilinear parabolic systems. III. Global existence. *Math. Z.*, 202(2):219–250, 1989.
- [2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.
- [3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [4] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.
- [5] M. Burger, M. Di Francesco, P. Markowich, and M.-T. Wolfram. Mean field games with nonlinear mobilities in pedestrian dynamics. *arXiv preprint arXiv:1304.5201*, 2013.
- [6] M. Burger, M. Di Francesco, J.-F. Pietschmann, and B. Schlake. Nonlinear cross-diffusion with size exclusion. *SIAM J. Math. Anal.*, 42(6):2842–2871, 2010.
- [7] P. Cardaliaguet, G. Carlier, and B. Nazaret. Geodesics for a class of distances in the space of probability measures. *Calc. Var. Partial Differential Equations*, 48(3-4):395–420, 2013.
- [8] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4):1273–1309, 2010.
- [9] J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.
- [10] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Differential Equations*, 224(1):39–59, 2006.
- [11] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.
- [12] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

- [13] L. Giacomelli and F. Otto. Variational formulation for the lubrication approximation of the Hele-Shaw flow. *Calc. Var. Partial Differential Equations*, 13(3):377–403, 2001.
- [14] U. Gianazza, G. Savaré, and G. Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Ration. Mech. Anal.*, 194(1):133–220, 2009.
- [15] A. Glitzky and A. Mielke. A gradient structure for systems coupling reaction-diffusion effects in bulk and interfaces. *Z. Angew. Math. Phys.*, 64(1):29–52, 2013.
- [16] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1981.
- [17] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [18] A. Jüngel. The boundedness-by-entropy principle for cross-diffusion systems. *arXiv preprint arXiv:1403.5419*, 2014.
- [19] A. Jüngel and I. V. Stelzer. Entropy structure of a cross-diffusion tumor-growth model. *Math. Models Methods Appl. Sci.*, 22(7):1250009, 26, 2012.
- [20] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction–diffusion systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(2005):20120346, 2013.
- [21] S. Lisini and A. Marigonda. On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals. *Manuscripta Math.*, 133(1-2):197–224, 2010.
- [22] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.
- [23] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352–1397, 2009.
- [24] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [25] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.
- [26] A. Mielke. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discrete Contin. Dyn. Syst. Ser. S*, 6(2):479–499, 2013.
- [27] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
- [28] F. Otto and M. Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255 (electronic), 2005.
- [29] R. Rossi and G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003.
- [30] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

ZENTRUM FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN, 85747 GARCHING, GERMANY

E-mail address: zinsl@ma.tum.de

E-mail address: matthes@ma.tum.de